

**Enhancing the cybersecurity for financial transactions using AI, ML and DL approaches**

A dissertation submitted as part of the requirements

for the degree of MSc. Computer science

at the school of computing and Built Environment

Birmingham city University, Birmingham, UK

Student: Chakrapani katam-23233635

January 2024

Supervisor: Dr. Debashish Das

Table of Contents

Contents

[Acknowledgements 4](#_Toc186979190)

[Abstract 6](#_Toc186979191)

[1. Introduction to the Research 7](#_Toc186979192)

[1.1 Background Research 7](#_Toc186979193)

[Figure-1: How AI and ML in cybersecurity 8](#_Toc186979194)

[1.2 Problem statement 9](#_Toc186979195)

[**Figure 2**, Loses to online payment fraud in 2020-2023 source: Statista.com 10](#_Toc186979196)

[Figure 3: It is related to the US global card fraud detection 11](#_Toc186979197)

[1.2.1 Research Problem 11](#_Toc186979198)

[1.3 Rationale 12](#_Toc186979199)

[1.4 Scope 12](#_Toc186979200)

[1.4.1 Aim 13](#_Toc186979201)

[1.4.2 Objectives 13](#_Toc186979202)

[1.4.3 Research Questions 14](#_Toc186979203)

[1.4.4 Nature of challenges 14](#_Toc186979204)

[1.5 Overview of the research 14](#_Toc186979205)

[2. Literature Review 15](#_Toc186979206)

[2.1 Introduction 15](#_Toc186979207)

[2.2 Domain Research 16](#_Toc186979208)

[2.2.1 Data Analytics in Financial Services 16](#_Toc186979209)

[2.2.2 Application of Deep learning and Machine Learning in predicting financial transactions 18](#_Toc186979210)

[2.2.3 cybersecurity in mobile financial transactions 20](#_Toc186979211)

[2.2.4 Integration of IOT in fraud financial transactions 21](#_Toc186979212)

[Research Gap Analysis Table 23](#_Toc186979213)

[TABLE-1 23](#_Toc186979214)

[3 Research Methodology 31](#_Toc186979215)

[3.1 Introduction 31](#_Toc186979216)

[3.2 Methodology 31](#_Toc186979217)

[Figure-4: Artificial Intelligence Automated threat detection in cybersecurity 32](#_Toc186979218)

[3.2.1 Research Design 32](#_Toc186979219)

[Table:2 33](#_Toc186979220)

[Figure- 5: Random Forest image 36](#_Toc186979221)

[Figure-6: Decision Tree image 37](#_Toc186979222)

[Figure-7: Artificial Neural Network image 38](#_Toc186979223)

[Figure-8: Artificial Neural Network inside the image 39](#_Toc186979224)

[Figure-9: Genetic algorithm with Neural Network (GANN) 40](#_Toc186979225)

[Figure-10: Confusion matrix 43](#_Toc186979226)

[Figure-11: ROC curve image 44](#_Toc186979227)

[Figure-11: pre-processing of the data 45](#_Toc186979228)

[4Experiments 46](#_Toc186979229)

[4.1 Data understanding 46](#_Toc186979230)

[Figure-12: Information of dataset 47](#_Toc186979231)

[4.2 Data preparation 47](#_Toc186979232)

[4.2.1 Data Pre-processing 47](#_Toc186979233)

[Figure:13: image illustrates the dataset there is no missing values 48](#_Toc186979234)

[4.2.2 Data Visualisation 48](#_Toc186979235)

[**Figure:14 Correlation Plot** 49](#_Toc186979236)

[Figure:15: Pie chart 50](#_Toc186979237)

[Figure:16: Heatmap of the data 51](#_Toc186979238)

[Figure:17: Count of different types of transactions 52](#_Toc186979239)

[Figure:18: Pie chart shows each type of transactions 53](#_Toc186979240)

[Figure:19: Target column is Fraud and Non-Fraud 54](#_Toc186979241)

[Figure:20: Histograms study each feature in my dataset 55](#_Toc186979242)

[4.2.3 Data Transformation 55](#_Toc186979243)

[Figure:21: Pie chart of the fraud transactions and Non fraud transactions in the data 57](#_Toc186979244)

[Figure:22: When I have applied SMOTE 58](#_Toc186979245)

[4.3 Model training 58](#_Toc186979246)

[4.4 Model Evaluation 60](#_Toc186979247)

[Figure:23 Accuracy of Machine learning Models Training 60](#_Toc186979248)

[Figure:24: Test accuracy of Machine learning models 61](#_Toc186979249)

[Table:3 62](#_Toc186979250)

[Table:4 63](#_Toc186979251)

[Table:5 64](#_Toc186979252)

[5.0 Conclusion and future works 65](#_Toc186979253)

[5.1 Introduction 65](#_Toc186979254)

[5.2 Objective Achievement 65](#_Toc186979255)

[5.3 Limitations and Future works 66](#_Toc186979256)

[References 70](#_Toc186979257)

# Acknowledgements

I would like to sincerely thank you Dr. Debashish Das sir for his incredible support and guidance throughout this project. Over the 12 weeks, he held regular meetings and provided valuable feedback every feedback. His advice and encouragement were crucial in helping me improve my work. Dr. Das dedication and patience made a significant difference, and I am truly grateful for his time and effort. His support was a key factor in the successful completion of this project. Thank you Dr. Das sir for being such a wonderful mentor and guide.

# Abstract

This project focus on improving the security of financial transactions by manipulating advanced technologies such as Artificial Intelligence (AI), Machine Learning (ML), Deep learning (DL). The primary goal is to be detect and prevent fraudulent activities in financial systems. We use a synthetic dataset to simulate real-world financial transactions and apply various data pre-processing to ensure data quality.

Our approach includes implementing anomaly detection, spam detection, intrusion detection, botnet detection and malware classification. We implement different machine learning algorithms and deep learning models to identify and mitigate security threats. Techniques like power Transformer and standardScaler are used to pre-process the data enhancing model performance. The model is trained and evaluated using an 80:20 train test split ensuring robust validations. Despite the challenges such as imbalance, our methods demonstrate significant improvements in detecting fraudulent activities. This project highlights the potential of AI, ML and DL in creating more secure financial systems, ultimately contributing to safer and reliable financial transactions.

# 1. Introduction to the Research

## 1.1 Background Research

In today’s digital age, the financial sector faces an ever-increasing cyberattacks. As financial transactions become more computerized, the complexity and sophistication of cyber threats have been increased, present significant risks to both institutions and purchaser. Standard cybersecurity tools including rule-based systems and firewalls are often incomplete in addressing these evolving threats (Maryam, Sylvester etc.2024). This demands the adaption of more advanced and adaptive security solutions.

Artificial Intelligence (AI), deep learning (DL) and Machine learning (ML) have all grown more useful in the fight against cybercrime. The technologies make it possible to analyse high amount of information in real-time, spot trends, and detect irregularities that could show fraud or security breaches ( Hilal, Yawney etc 2022). AI-driven system would learn from historical data to predict and find out feature attacks, While ML algorithm can continuously improve their accuracy by adapting to new threat patterns (Ranjan, P. and Dahiya, S., 2021). DL, a subset of ML, leverages neural networks to process complex data structures, providing even greater precision in detection of threats and moderation.

A diagram of a network security system

Description automatically generated

# Figure-1: How AI and ML in cybersecurity

AI, DL, and ML integration in cybersecurity frameworks improves the capacity to quickly and efficiently identify and address cyberthreats. By lowering false positives, these technologies also increase threat detection accuracy. Thereby increasing the efficiency of security operations. By manipulating ML, AI and DL financial institutions can build more resilient cybersecurity infrastructures, ensuring the protection of important financial data and maintain customer trust (wang.s,Asif.M etc 2024).

This report explores the various ML, AI and DL approaches employed in enhancing cybersecurity for financial transactions. It examines the current state of these technologies, their applications in fraud detection and prevention, and the problems and prospects for the quickly developing field (Chatterjee.P,Das.D etc, 2024). By means of an extensive analysis of current literature reviews and case studies, this work aims to be provided a robust framework for implementing advanced cybersecurity measures in financial sector (Wang, S., Asif, M., Shahzad, M.F,2024) .

Cyberattacks target the banking sector because of the large amounts of information it can’t handles and the possible financial gains for attackers. Cyber threats in this sector range from identity theft and spam to sophisticated malware and ransomware attacks is crypto trojans ( Rana.M, shah.M etc,2024). Traditional cyber techniques for security features include signature-based detection systems and firewalls, have proven insufficient in addressing these evolving threats (Kinga.S,Mugerwa etc,2024). The growing complexity and frequency of cyberattacks necessitate the use of more sophisticated and flexible security measures.

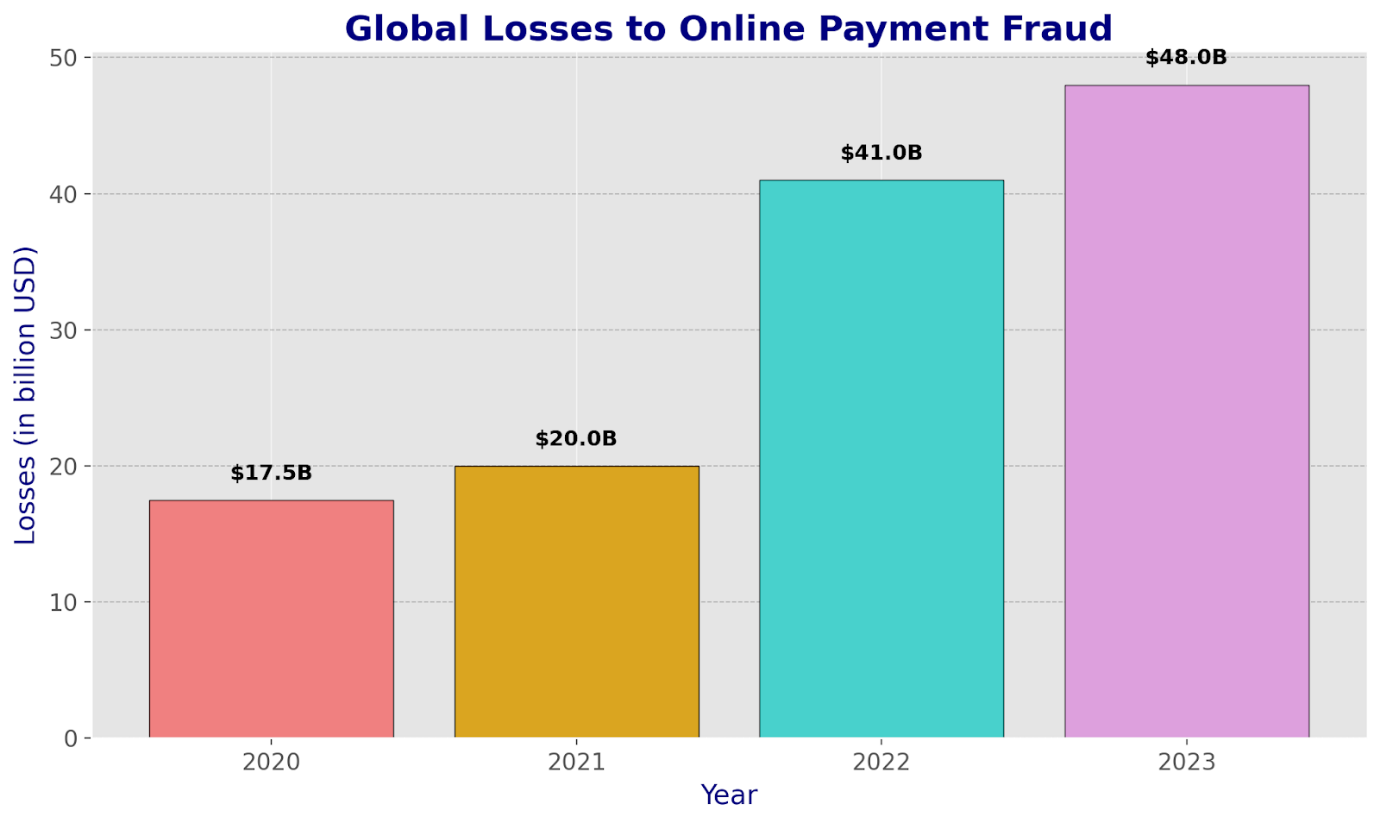
The ultimate goal is to in order to guarantee the safety of sensitive financial information, maintain consumer trust, as well as support the financial systems general stability.

# 1.2 Problem statement

The quick online transactions and services of the banking industry make it more susceptible to attacks by the cybercriminals. Conventional cybersecurity tools like signature based detection systems and firewalls, are often incomplete in addressing the sophisticated and evolving nature of these threats ( Luong.K,Barry.B etc,2024). Cybercriminals employ advanced techniques to breach security systems, leading to significant financial losses, data breaches and erosion of buyer trust ( Dakic.V, Regvart.D etc,2023).

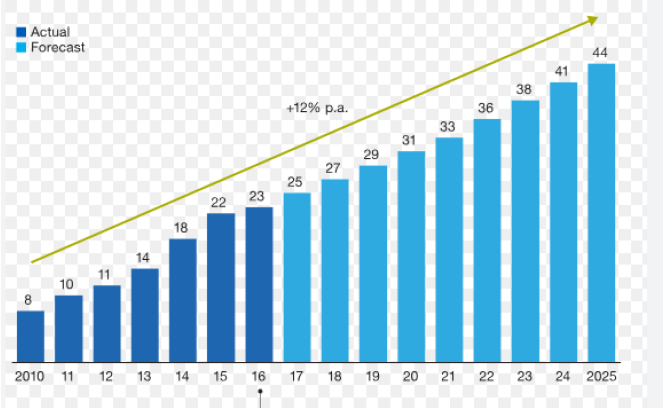
Despite the potential of Machine Learning (ML), Deep learning (DL) and Artificial Intelligence (AI) to enhance cybersecurity, their adoption in the financial sector faces several challenges ( Sarker, Maglaras etc,2024). These includes the complexity of integrating AI-driven solutions into existing systems, large datasets are required in order to train the models like Logistic Regression, Random Forest (RF), Support vector machines (SVM), Naïve Bayes and Decision tree. Deep learning models like Artificial Neural Network (ANN), Conventional Neural Network (CNN) and Long-short term memory (LSTM) and the difficulty of in interpreting and explaining AI decisions. Additionally, there is a lack of comprehensive frameworks that leverage ML, AI and DL to provide robust and adaptive cybersecurity measures tailored to the unique needs of financial transactions ( Kavitha.D and Thejas.S 2024).

The main of the study support the growth of more resilient cybersecurity infrastructures in the financial sector, untimely protecting important information financial data and maintaining customer trust.



# **Figure 2**, Loses to online payment fraud in 2020-2023 source: Statista.com

The online payments in the year resulted in losses for both banks and customers, as the accompanying data illustrates. Advanced tools are being used by cybercriminals to access the most recent world bank reports and financial activities. In 2020, fraudulent transactions of 17.5 billion in the world. The subsequent year, 2021, saw 20 billion dollars in online fraud transactions. The fraud transaction that occurred in 2022 was 50% more than the one that occurred the year before. 41 billion is what it means. The following year, 2023, saw 48 billion fraudulent transactions conducted online.



# Figure 3: It is related to the US global card fraud detection

In the above figure is related to united states of America of every year it increases the step-by-step year. When we are growing the financial transactions, fraud also increased. It figures graph represents global and domestic card and prepaid cards and online fraud transactions.

### 1.2.1 Research Problem

The financial sector is increasingly targeted by sophisticated cyber-attacks, resulting in significant financial losses, data breaches, and breakdown of customer trust ( Gulyas.O and Garbor kiss.G 2023). Conventional methods of cybersecurity, like detection systems and firewalls based on signatures are frequently unable to handle the dynamic nature of these threats, are often insufficient in dealing with these threat’s changing nature. As cybercriminals adopt advanced techniques, security solutions that are more resilient and flexible are desperately needed ( Mallick, M.A.I. and Nath, R., 2024).

Artificial Intelligence (AI), Deep learning (DL) and Machine Learning (ML) offer promising capabilities for enhancing cybersecurity. These systems examine massive volumes of real-time data, which also find trends and anomalies that could point to fraud or security breaches. However, the incorporating AI-powered solutions into already-existing financial systems presents several challenges. This includes complexity of implementation, large, high-quality datasets are necessary for model training, and the challenge is interpreting and explaining AI decisions (Ding.W & Hawash.H 2022).

This research problem addressed in this dissertation is to explore and evaluate various ML, AI and DL approaches to enhance cybersecurity in financial transactions. The goal is to create a through structure that improves the scalability and interpretability of AI-driven solutions. The goal of this research is to help the financial industry build more robust cybersecurity infrastructures.

## 1.3 Rationale

According to the problem statement in section 1.2, The financial transactions increasingly move to digital platforms, the sector becomes more susceptible to sophisticated cyber-attacks. In ML and DL that is used to find out the fraud and is fraud classifying the tasks 0 means not fraud and another type 1 is Fraud. These attacks not only cause significant financial losses, but also destroy private information and customer confidence. The dynamic and intricate nature of these threats frequently makes traditional cybersecurity techniques inadequate.

Financial institutions that use AI,DL and ML may create more resilient and flexible security frameworks that are better equipped to handle new threats. The rational for this study aims to investigate how AI, DL and ML can be integrated into cybersecurity strategies to enhance the protection of financial transactions.

## 1.4 Scope

Cybersecurity threats in the financial sector an in-depth analysis of the types of cyber threats faced by financial institutions, including phishing, malware, ransomware like crypto trojans and insider threats. This section will also cover the impact of these threats on financial transactions and data security.

A detailed examination of the various ML, AI and DL techniques used in cybersecurity. This includes supervised and unsupervised learning algorithms, neural networks and anomaly detection methods. This study will explore how these technologies can be applied to detect and find out cyber threats in real-time.

The development of a framework for integrating AI, DL and ML into existing cybersecurity infrastructures. This will include the design of models, techniques for gathering and pre-processing data as well as the implementation of AI-powered security solutions. The effectiveness of AI, DL and ML approaches in enhancing cybersecurity. This will involve evaluating the accuracy, efficiency and scalability of the proposed solutions through experiments and case studies.

Identification and analysis of the challenges associated with implementing ML, AI and DL in cybersecurity. This provides information related to data quality, model interpretability and the integration of AI technologies into legacy systems. Examine new developments and potential lines of inquiry in the area of AI-driven cybersecurity.

### 1.4.1 Aim

This research aims to predict fraud transactions from datasets obtained which will help to provide an in-depth understanding of the types of cyber threats faced by the financial sector as well as the shortcomings of conventional cyber security methods. The goal is to help create cybersecurity that is more resilient and adaptable measures in the financial sector. To evaluate the performance, accuracy and scalability of the proposed AI-driven cybersecurity solutions through experiments.

### 1.4.2 Objectives

The illustrate describes the research objectives:

* Analyse and identify possible cybersecurity risks aimed at financial transactions using AI, ML and DL.
* Improve real-time cyber threat identification and prevention by creating sophisticated algorithms.
* AI-driven solutions were implemented to minimize the effect of cybersecurity incidents by automating responses.
* To improve the security and integrity of sensitive financial information, apply machine learning and deep learning techniques.
* Optimize AI and ML models to minimize false positives in threat detection.
* Implement AI-based methods to minimize false positives in threat detection.

# 1.4.3 Research Questions

The following describes the research questions:

1. Which AI, ML and Dl methods are best for identifying and stopping cyberattacks in financial transactions?
2. How can AI, ML and DL technologies be integrated into existing cybersecurity frameworks within financial institutions?
3. What challenges do financial institutions face when implementing AI, ML and DL for cybersecurity, and how can these challenges be mitigated?

### 1.4.4 Nature of challenges

* Labelled datasets of good quality are necessary for training AI, DL and ML models. However, obtaining such data. Due to privacy issues and the sensitive nature of financial information, data can be challenging. Additionally, the data can be comprehensive and representative of various cyber-attacks to ensure the models are effective.
* ML, AI and DL models, especially deep learning models in particular can be intricate and challenging to understand. Financial institutions need to understand how these models make decisions to ensure they are reliable and to comply with regulatory needs. Lack of interpretability can be hinder trust and adoption of these technologies.
* DL, ML and AI models need to process large amounts of data instantly to be effective in cybersecurity. One major difficulty is making sure these models can scale to accommodate the data load without sacrificing speed.

## 1.5 Overview of the research

The financial transactions and services are becoming more and more digitalization, the financial sector is becoming more and more susceptible to sophisticated cyberattacks. Conventional cybersecurity methods are frequently insufficient in addressing these evolving threats, necessitating the adoption of more advanced and adaptive security solutions. This research explores the potential of Artificial Intelligence, Deep Learning and Machine Learning to enhance cybersecurity in financial transactions.

**Current cybersecurity Landscape:**

To begin, Analyzing the types of cyber threats faced by financial institutions, such as phishing, malware and ransomware. Assessing the limitations of traditional cybersecurity measures in mitigating these threats.

**AL, ML and DL Techniques:**

Evaluating various AI, ML and DL algorithms like Decision Tree, Logistic Regression, Naïve Bayes, Random Forest classifier, Support vector Machines in many more algorithms in machine learning. In deep learning models like Conventional Neural Network, Artificial Neural Network (ANN), Long-short term memory (LSTM). This algorithms and models can be applied to cybersecurity.

Determine the best methods for immediately identifying and stopping cyberattacks, such as financial activities. Algorithms find is fraud or non-fraud transactions.

**Implementation Framework:**

Furthermore, Developing a comprehensive framework for integrating AI, ML and DL technologies into existing cybersecurity infrastructures. Outlining the steps for gathering data, preparing it, training models and deploying them.

**Challenges and Future directions:**

Identifying the challenges associated with implementing ML, AI and DL in cybersecurity such as model interpretability, data quality and legacy system integration. Investigating new trends and possible advancements in cybersecurity powered by artificial intelligence.

# 2. Literature Review

## 2.1 Introduction

The financial industry is a major target for attackers because of the high value of the information it manages and the possible financial advantages for them. As financial transactions become increasingly digitized, the cyber dangers have become more sophisticated and complex, posing significant risks to both institutions and consumers. Traditional cybersecurity actions like rule-based systems and firewalls are often inadequate in addressing these evolving threats. This necessitates the adoption of more advanced and solutions for adaptive security.

## 2.2 Domain Research

### 2.2.1 Data Analytics in Financial Services

Data analytics is the methodical examination of few datasets to identify significant patterns, trends and insights in the financial services industry. This technique enables experts to make well-informed choices maximise and improve overall performances.

1. **Fraud Detection and Prevention:**

In order to identify irregularities and possibly fraudulent activity in real-time, advanced analytics approaches like machine learning algorithms can examine transaction patterns (Hilal.W & Gadseden. S.A 2022). This proactive strategy aids in stopping fraud before it starts.

A large bank analyses transaction patterns and looks for anomalies that might point to fraudulent activity using machine learning techniques (Gandhar. A Pandey.A.K etc). By monitoring transactions in actual time, the bank can quickly identify and block suspicious activities, reducing the risk of fraud.

1. **Risk Management:**

Financial organisations use data analytics to assess and mange a range of hazards, such as market, operational and credit risks. Institutions can anticipate possible hazards and implement preventive actions by looking at market patterns and historical data (Olisakwe.HC & Esiri.AE etc 2024).

1. **Customer insights and Personalization:**

Financial companies can better understand the behaviour and preferences of their customers by using data analytics. Using this information, products and service might be tailored to each person’s needs, boosting customer satisfaction and loyalty ( Yum k & Kim .J 2024).

1. **Operational Efficiency:**

Through operational data analysis, financial organisations can pinpoint inefficiencies and streamline procedures. Cost reductions, increased productivity and better resource allocation result from this.

A major bank implemented an automated loan processing system to enhance its operational efficiency (Bueno, L.A., Sigahi, T.F. etc 2024) Traditionally, the loan approval process involved multiple manual steps, including data entry, document verification and credit assessment, which were time-consuming and prone to errors (Umar, M.A. & Lano, K., 2024).

1. **Predictive Analytics:**

Predictive models estimate future patterns based on historical data and behaviours. In the financial sector, this would be applied to predict market moments, customer churn and loan defaults, enabling institutions to make data-driven decisions ( Rajasekhar.E, Balachandran.S etc 2024).

Predictive analytics is used by a financial institution to evaluate loan applicant’s creditworthiness. By analysing historical data such as credit scores, transaction histories and employment records the institutions machine learning models predict the likelihood of an applicant defaulting on a loan (Twala.B & Pretorius.J etc 2024). This allows the institutions to make informed lending decisions reduce the risk of defaults and optimize its loan portfolio.

1. **Credit Scoring and Underwriting:**

By adding a greater variety of data points to credit scoring models, analytics can improve them and produce more precise evaluations of creditworthiness. For instance, past transactions and social media usage. This result in improved risk management and more precise evaluations of creditworthiness ( Bravo, C., Ríos, S.A. etc).

1. **Investment Strategies:**

Financial analysts use data analytics to develop and refine invest strategies. By analysing market data and trends, they can identify new investment opportunities and optimize portfolio performances (Zhu, Q., Zhou, X. and Liu, S., 2023).

Financial organisations can enhance their operations by utilising data analytics not only increase the effectiveness of their operations and customer satisfaction but also enhance their ability to detect and prevent fraud, manage risks, and comply with regulatory requirements.

**Benefits of Data analytics in Financial Services:**

Data driven insights enable more accurate and timely decisions, reducing uncertainly and enhancing strategic planning.

Identifying inefficiencies and optimizing processes can lead to significant cost savings.

Personalized services and proactive engagement improve customer satisfaction and loyalty.

Streamlined operations and better resource management enhance overall efficiency (Gupta, A. and Agarwal, P., 2024,).

Leveraging data analytics provides a strategic edge in a highly competitive market.

### 2.2.2 Application of Deep learning and Machine Learning in predicting financial transactions

The financial services sector has seen a substantial transformation thanks to machine learning and deep learning. Which have made it possible to forecast financial activities more precisely and effectively:

* **Fraud Detection**:

**ML algorithms**: In order to identify fraudulent transactions,

Supervised learning algorithms like Logistic Regression, Random

Forest classifier, support vector machines (SVM) and DecisionTree

(DT) analyse past transactions data and look for patterns that point

To fraud. For instance: Paytm employs machine learning models to

Evaluate transactions data in real-time, identifying questionable

Behaviours and stopping fraudulent transactions.

**DL Models**: Large amount of transaction data can be processed by deep learning models neural networks, such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, to identify minute irregularities that might point to fraud (Govindarajan, M., Vijayakumar, V. etc 2023). For instance: JP Morgan chase employs DL models to enhance its fraud detection capabilities, reducing false positives and improving detection accuracy.

* **Credit Scoring:**

Machine learning techniques and models analyze numerous data points, including income, credit history and spending behaviour to predict an individual’s creditworthiness ( Mienye,E. ,Jere.N etc 2024). These models can be providing more accurate credit scores compared to traditional methods. For example: FICO uses ML algorithms to enhance its credit scoring system providing more precise assessments of credit risk.

DL approaches and models can handle complex, non-linear relationships in the data improving the accuracy of credit scoring systems (Gambacorta, L., Huang, Y. 2024). For example: The majority of business including Bajaj Finance, employ DL to examine information from sources such as social media activity and transaction history to assess credit risk for people with little credit history.

**Algorithmic Trading:**

To forecast future price fluctuations and automatically execute trades, machine learning models and algorithms examine past market data ( Parente.M & Trerotola.M etc 2024). By adjusting to shifting market conditions, these models help improve trading. For example: Renaissance Technologies, a hedge fund, uses ML models to develop trading algorithms that have consistency outperformed the market.

DL techniques and models such as Recurrent Neural Network (RNN) and Transformers analyse time-series data to forecast market patterns and stock prices. For instance: Quantitative trading firms like two sigma’s use DL models to process vast amounts of financial data and make informed trading decisions.

* **Risk Management:**

Models and uses of machine learning using past data and market patterns, assess and predict a range of risks, including operational, market and credit risk (S., Mohan, L. and Reddy, Y.R., 2020). For example: Goldman Sachs uses ML models to predict market vitality and manage risk exposure.

Financial institutions can more successfully reduce risks by using deep learning models, which can analyse enormous volumes of data to detect possible threats and deliver real-time risk evaluations (Shah, V., 2021). For instance: BlackRock, the biggest asset manager in the world, uses DL models to improve its risk management techniques.

* **Customer insights and personalization:**

Financial organisations can be provided individualised goods and services by using machine learning algorithms to examine consumer data and find trends and preferences. For example: Lloyd’s bank uses ML to analyze customer interactions and provide personalized financial advice through its virtual assistant, United kingdom.

Deep learning models can better understand customer behaviour and enhance personalisation efforts by analysing unstructured data, such as social media activity and customer reviews. For example: companies like capital one use DL to enhance their customer experience by tailoring services to individual needs.

Financial organisations can improve their decision-making processes, increase their predictions capacities and offer their clients better services by utilising ML and DL. These technologies only increase efficiency but also help in making risks and detecting fraud more effectively.

### 2.2.3 cybersecurity in mobile financial transactions

Cybersecurity in mobile financial transactions is of paramount importance due to the increasing reliance on mobile devices for banking, payments and other financial activities. The primary threats include phishing attacks, where fraudsters attempt to steal sensitive information by masquerading as trustworthy entities and malware, which is malicious software designed to infiltrate mobile devices and steal data or cause damage.

Man-in-the-middle attacks, where communication between two paties is intercepted and altered without their knowledge and data breaches, which involve unauthorized access to sensitive financial data are also significant concerns. To combat these threats, several security measures are essential. Encryption ensures that data is protected during transmission and storage, making it inaccessible to unauthorized users.

Multi-factor authentication (MFA) adds an extra layer of security by requiring multiple forms of verification, such as passwords and biometric data. Regular software updates are crucial to protect against known vulnerabilities, and user education is vital in teaching users about safe practices, such as recognizing phishing attempts and using secure networks.

Looking to the future, artificial intelligence (AI) can be utilized to detect and respond to threats in real-time, significantly enhancing the security of financial transactions. Blockchain technology offers secure and transparent transactions, reducing the risk of fraud. Additionally, regulatory compliance with standards like GDPR and PSD2 ensures the protection of user data and privacy.

In conclusion, enhancing cybersecurity in mobile transactions involves a combination of advanced technologies. Robust security measures, and user awareness. By addressing these areas, financial institutions can better protect their customers and maintain trust in their services. The integration of AI and blockchain technology along with strict adherence to regulatory standards, will play a crucial role in the future of mobile financial security.

### 2.2.4 Integration of IOT in fraud financial transactions

The integration of the internet of things (IOT) in combating fraud in financial transactions is revolutionizing the banking and financial sectors. Smart sensors and other connected devices are examples of IOT devices that provide real-time monitoring and data collecting, both of which are essential for identifying fraudulent activity. Financial institutions can collect enormous volumes of data from several sources, such as transaction patterns, user behaviour, and environmental conditions, by utilizing IOT.

After that, this data is examined in real-time for anomalies and possible fraud using sophisticated algorithms and machine learning models. For instance, Devices enabled by the internet of things can keep an eye on transactions for odd trends, such as several transactions from several locations in a little amount of time, which could point to fraud.

Additionally, IoT enhances security through biometric authentication and blockchain integration, ensuring that transactions are secure and tamper-proof. Continuous network monitoring by IoT devices helps in identifying and mitigating potential threats before they can cause significant damage. The application of IoT to fraud detection enhances financial institutions overall security infrastructure in addition to increasing the precision and spedd of detecting fraudulent transactions.

As IoT technology continues to evolve, its integration with artificial intelligence and machine learning will further strengthen the capabilities of fraud detection systems making financial transactions safer and more reliable.



# Research Gap Analysis Table

## TABLE-1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **S.NO** | **Article** | **Author and year** | **Journal** | **Model/Techniques** | **Findings** | **Limitations** |
| 1. | Cyber defense for financial fraud detection applications in the era of artificial intelligence and machine learning | B.Narsimha, V.Raghavendran, Rajyalakshmi,kasi Reddy Bhargavi, Naresh. (2022) | IJEER | Random Forest,  Decision trees,  Neural network, KNN,  Naïve Bayes. | The Random forest model was the best accuracy 82.94% | In this article, the author explained they applied models in the fraud detection using AI, Ml and DL approaches’ fraud detection rate :38.62  Naïve bayes: 38.46 Random Forest:42.65. |
| 2. | A thorough framework for enhancing financial cybersecurity in the United States: using AI and machine learning into fraud detection systems. | Adijat Bello, Folorunso,jane, Eunic Ejiofor(2023) | European Journal of Computer Science and Information Technology | Logistic Regression, decision trees, random Forest, support vector machines and Neural Networks. | Through real-time analysis of massive datasets, pattern recognition and fraud activity prediction, machine learning and artificial intelligence significantly increase the accuracy and speed of fraud detection. | In this article, They have explained Ai in fraud detection in ML and Deep learning.  Financial institutions need to invest in technology infrastructure and skilled personnel. |
| 3. | Using real-time machine learning for financial security using artificial intelligence in fraud detection. | Ravi Teja Potla(2023) | JAIRA | Random Forest, Gradient Boosting Machines, support vector machines, decision trees | AI and ML can scale to handleincreasing volumes of transactions and data making them suitable for large financial institutions. | Poor or biased data can undermine model accuracy and effectiveness. |
| 4. | Examining how AI-based cyber security affects management of the Financial sector. | Shailendra Mishra(2023) | MDPI | KNN, Decision trees, random forest, ANN | AI algorithm significantly improve the detection of cyber threats by analysing large volumes of data in real-time, identifying patterns and potential attacks. | AI systems can sometimes generate false positives or false negatives. Which can lead to inefficient and missed attacks. |
| 5. | Blockchain, deep learning , Machine learning and artificial intelligence in baking and financial services. | Mallikarjuns paramesha, Nitin Liladhar Rane, Jayesh  (2024) | Partners universal multidisciplinary Research journal  (PUMRJ) | Neural network, support vector machines, decision trees and logistic regression | In this research paper, they explained This models are good for financial fraud detection. | While these technologies are scalable smaller financial institutions might struggle to leverage them fully due to resource constraints. |
| 6. | Cybersecurity of online Financial systems using Machine learning technique. | Sorin-Ionut,  Stefania-Loredana Nita  (2024) | IEEE | Support vector machines, Artificial neural network, decision trees, Deep learning and Random Forest | In this paper, they have explained in the F1 score validation and gini impurity is good 0.87,0.88,,0.88 and gini impurity 0.81,0.82,0.82 | The proposal model presenting better performances in fraudulent transactions detection balance 23.54% or imbalanced 20.37% |
| 7. | Improving machine learning and imbalance mitigation for financial transaction fraud detection. | Mohammad, Rama Khaled, Fakhri Alam Khan  (2024) |  | Bayesian algorithm, K Nearest Neighbour, support vector Machines, decision trees | Deep learning 53.4% accuracy.  Random forest is recall rate of 0.80% precision rate 0.91% and F-score 0.85% | Using large datasets for training ML models, making them less effective. For instance, they might introduce subtle changes to transaction data to bypass detection. |
| 8. | Machine Learning approaches for Enhancing Fraud prevention in Financial Transactions | Bello, Oluwabusayo Adijat and Folorunso, Adebola.  (2024) | Publication of the European Centre for Research Training and Development | logistic regression, decision trees, and neural networks, support vector machines, Random Forest, Naïve bayes, logistic regression | By analyzing sequential data and identifying complex patterns over time, CNN and RNNs can improve the detection of complex fraud schemes. | Fraudsters can use adversarial techniques to manipulate ML models, making them less effective. For instance: they might introduce subtle changes to transaction data to bypass detection. |
| 9. | Using Machine learning and deep learning to strengthen Bangladesh’s Financial infrastructure for Banking cybersecurity | Badiuzzaman, Tonmoy and Farabi  (2024) | springer | logistic regression, support vector machines, Random forest Conventinal neural network(CNN) and decision trees | These technologies enhance data privacy and security by employing advanced encryption methods and secure data handling practices. | Developing and maintaining ML and DL systems can be complex and expensive. This includes costs related to hardware , software and skilled |
| 10. | A systamatic review for preventing prospective solutions in financial transactions. | Mahfujur Rahman Faraji, Fisan shikder, Hasan, Islam  (2024) | International journal of Religion | Random Forest, k-Nearest Neighbours (KNN), support vector machines (SVM), Decision tree, Multinomial Naïve Bayes, logistic regression | AI algorithms are constantly learning from fresh data at spotting new fraud trends and online dangers. | Effective model training requires high quality labelled datasets, but acquiring such datasets can be difficult because of privacy concerns and the dynamic nature of financial transactions. |
| 11. | Enhancing cyber security by predicting malwares using supervised machine learning models | Dinesh kalla, samaah and sivaraju kuraku    (2021) | International journal of computing and Artificial Intelligence | Support vector machines, Random forest, decision trees, Logistic regression, KNN | Random forest performed the best with 99% accuracy, and 0.99 precision, 1.00 recall and a 0.99 F1 score  And K-nearest neighbour accuracy is 97% and support vector machines reaches 75% accuracy and decision trees 98% accuracy | These algorithm can analyse vast amounts of data to identify patterns and that might indicates a security beach |
| 12. | Machine learning and Deep learning approaches for cybersecurity | Asmaa Halbouni, Surya, Mohamed  (2022) | IEEE |  | DL techniques, including transformers and graph neural network (GNN’s) are used to analyse the vast amount of threats intelligence data, identifying relationships and predicting potential threats. | DL models, especially deep neural networks, require significant computational power and resources which can be costly. |
| 13. | The role of deep learning in ensuring privacy integrity and security: applications in AI-driven cybersecurity solutions | Joseph, Moshood yussuf, okusi, Tempitope | WJARR | Convolutional Neural Network, Long short Term Memory, | DL models can incorporate different privacy techniques to ensure that individual data points cannot be distinguished. | DL models require vast amounts of labelled data to train effectively . In cybersecurity, obtaining such data can be challenging due to privacy concerns and dynamic nature of threats. |
| 14. | Optimizing fraud detection in financial transactions with machine learning and imbalance mitigation | Ezaz Mohammad Al-dahasi, Rama Khaled, Alam khan, Gwanggil Jeon (2024) |  | Bayesian algorithms, K Nearest Neighbour, support vector machines, decision tress, Long-short term memory (LSTM) | LSTM and Gated Recurrent unit (GRU) 94.98% and 94.62% Logistic regression 99.9% support vector machines 99.9% Gaussian NB 99.9% Random forest classifier 100% | Fraud patterns evolve over time, and ML models need to be continuously updated and retained to adapt to new types of fraud. This ongoing maintains can be resource intensive. |
| 15. | Deep learning and Artificial Intelligence Framework to Improve the cybersecurity | D Ghallani (2022) | Researchgate.net | Conventional neural network, recurrent neural network. | CNN’s high accuracy in detecting network intrusions by learning complex patterns in network traffic data. They are particularly effective in identifying that traditional methods might miss. | Deep learning models can overfit to training the data, performing well on known data but poorly on new, unseen data. This can be limit their effectiveness in detecting novel threats. |
| 16. | A Deep learning Approach for cyber security and Financial Fraud detection and classification | Uday kumar, Joshi, Boomiga, Sugumar (2023) |  | Decision Trees, Naïve Bayes, Support vector machines, Random Forest, XG Boost, Long-short term memory (LSTM), CNN | Support vector machines precision is 81.88% and accuracy is 82.02% Random forest precision is 81.09% accuracy is 82.66% CNN precision is 77.08% and accuracy is 76.10% | Cyber threats and fraud patterns evolve rapidly, requiring deep learning models to be continuously updated and retained to adapt to new types of threats. |
| 17. | State of the art, challenges and future directions | Sema Admass, Yayeh, Diro (2024) | ScienceDirect | LSTM, Decision trees, Naïve Bayes, Support vector Machines (SVM) | Integrating advanced machine learning techniques and developing more comprehensive evaluation metrices. | Many tracking algorithms struggle with high-density scenes and complex interactions between objects. Additionally, the computational cost can be high, making real-time tracking difficult. |
| 18. | Deep learning Algorithms for cybersecurity applications: A Technological and status Review | Priyanka Dixit, Sanjay silakari (2021) | ScienceDirect | Recurrent Neural Network (RNN), Long short Term memory (LSTM) | These models have shown accuracy rates of around 90% to 98% in detecting phising attacks by analyzing url’s and email content. | While deep learning models can be highly effective, scaling them to handle large, real -world datasets diverse cyber threats remains a challenge. |
| 19. | Using Deep learning to detect anomalies in Financial transactions with SAP | Surya sai Ram Parimi (2024) | International Research Journal | Recurrent Neural Network (RNN), convolutional Neural Networks (CNNs), Long short term memory (LSTM) | Autoencoders, Recurrent Neural Network and Convolutional Neural Network are commonly used for anamoly detection in financial transactions. These model shows significant improvent in detecting anomalies compared to traditional methods. | Achieving real-time anomaly detection is challenging due to high computational demands of deep learning models. This can be a critical limitions in scenarios where immediate action is required to prevent fraud. |
| 20. | An Intelligent cyber security phishing detection systems using deep learning techniques | Mughaid, Shadi AlZU’bi, Esraa Abu Elsoud (2022) | Springer Nature Link | Support vector machines, decision tree, logistic regression, Neural network, | SVM accuracy is 0.83% and boosted decision tree is 88% logistic regression is 81.4% and Neural Network is 80.66% and decision forest is 86.9%. | Training and deploying deep learning models require substantial computational power and resources, which can be a barrier for real-time phising detection systems. |
| 21. | A thorough analysis from the perspective of neural networks and Deep learning | Iqbal H. Sarker  (2021) | Spring Nature | Aritificial Neural Network (ANN), Multi-layer perceptron, convolutional neural Network(CNN), Recureent neural network or Long short-term memory. | Common deep learning models include Convolution Neural Network for image preprocessing , recurrent Neural Network for sequential data and Generative adversarial Networks for generating new data. | Many real-world datasets are imbalanced with a disproportionate number of example in different classes which can affect model performance and bias the results. |
| 22. | Application of deep learning to cybersecurity | Samanesh Mahadavifar, Ali Ghorbani (2019) | ScienceDirect | Convolutional Neural Network(CNN), Recurrent Neural Network(RNN), Long-short term Memory(LSTM) | Deep learning approaches, including Long-short term memory (LSTM) networks, are used to detect phishing attacks by using analyzing URL’s, email content and other features. | Cybersecurity datasets often have imbalanced classes, with far fever examples of attacks compared to normal behavior, which can affect model performances. |
| 23. | Deep learning techniques and IEEE standards are used to protect internet of things devices from cyberattacks. | Nayem Uddin Prince, Mohd Abdullah AI mamum, Ahmed Olabisi, Obyed Ullah Khan, Adedokun Bidemi Akeem, Abuh Ibrahim sani (2024) | IEEE | Convolutional Neural Network, Long short-term Memory (LSTM) | Achieving real-time detection is challenging due to high computational demands of deep learning models. | Deep learnings models can be highly effective scaling them to handle large, real-world datasets and diverse cyber threats remain a challenge. Effectively managing and processing large volumes of data in crucial for practical deployment. |
| 24. | Methods strategies applications difficulties and prospects for further study in machine learning and deep learning. | Dimple Patil, Nitin Liladhar, Pravin Desai, Jayesh Rane (2024) | DeepScience | Long short term memory (LSTM), Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) | Difficult in scaling models to handle large datasets on real-time processing . | Deep learning models, particularly those using architectures like CNNs and transformers have significantly improved accuracy in tasks such as image recognition and natural language processing. |
| 25. | Deep learning Methods and IEEE standards to protect Internet of Things (IoT) Devices from cyberattacks. | Nayem Uddin prince, Mohd Abdullah AI Mamun, Ahmed Olabisi Olajide, obyed Ullah khan, Adedokun Bidemi Akeem, Abuh Ibrahim sani (2024) | IEEE | Long-short term memory, Convolutional Neural Network, Recurrent Neural Network | Deploying deep learning models in real-time IoT security raises concerns about data privacy and the ethical implications of minoriting and analyzing user behaviour. | Deep learning models such as CNNs and LSTMs have shown high effectiviness in detecting anomalies and intrusion in IoT networks. These models can alayze vast amounts of data to identify patterns indicative of cyber threats. |

# 3 Research Methodology

## 3.1 Introduction

Any academic study or scientific investigation must include a research technique since it offers a methodical strategy and plan for addressing the research questions or hypotheses. The research outlines the specific techniques tools such as the procedures used to collect analyse and interpret data to ensure that it is conducted in a systematic and reproducible manner. A clear research approach improves the validity and dependability of the study’s conclusions and enables a thorough comprehension of the research subject.

In this report, the research methodology is designed to explore the application of Artificial intelligence (AI), Machine Learning (ML) and Deep learning (DL) in enhancing cybersecurity for financial transactions (Mishra S 2023). The chosen methodology combines both qualitative and quantities approaches to provide a holistic view of the subject matter. This mixed methods approach allows for the collection of diverse data types, facilitating a more robust analysis and interpretation of the results.

## 3.2 Methodology

The methodology outlines the systematic approach and procedures used to conduct the research, making certain that the study is well-organised, repeatable and able to successfully answer the following questions.

A diagram of a process

Description automatically generated

# Figure-4: Artificial Intelligence Automated threat detection in cybersecurity

### 3.2.1 Research Design

The study uses a mixed-approaches strategy that incorporates both quantitively and qualitative methods. With the strategy the advantage of both approaches is combined to provide a more thorough examination.

.

* **Qualitative Methods:**

These includes interviews and case studies to gain in-depth information about the present situation of cybersecurity in financial institutions and the challenges faced by in implementing AI, ML and DL technologies ( Parkinson, S. and Khan, S., 2024).

* **Quantitative Methods:**

These include assessing how well AI, ML and DL techniques identify and stop cyberthreats using statistical analysis and machine learning models (Geetha, R. and Thilagam, T., 2021).

**Data collection**

A crucial part of the research process is data collecting, which includes obtaining both primary data:

The aim of the project on the based on the assumptions and understanding of the overall the dataset structure such as variables, quality of the data is evaluated.

The dataset used for financial transactions model. In the below the dataset describes below:

# Table:2

|  |  |  |
| --- | --- | --- |
| **Attribute name** | **Data Type** | **Description** |
| Step | Numeric | One real-world hour is represented by this unit of time. There are 744 steps in all, which is equivalent to a 30-day simulation. |
| type | Categorical | The nature of the deal.CASH\_IN, CASH\_OUT,DEBIT,PAYMENT,AND TRANSFER are the available forms. |
| amount | Numeric | The transaction’s value in local currency. |
| nameOrig | Categorical | The consumer who made the buy. |
| oldbalanceOrg | Numeric | The customer’s starting balance prior to the transactions |
| newbalanceOrig | Numeric | The buyer new balance following the transactions |
| nameDest | Categorical | The transaction's receipt |
| oldbalanceDest | Numeric | before the transaction, the receipt's initial balance.No information is available for clients whose names begin with "M." |
| newbalanceDest | Numeric | following the transaction, the recipient's new balance. like "oldbalanceDest."Customers with names that begin with "M" are not given any information. |
| isFraud | Numeric | Indicates whether a transaction is fraudlent . in this dataset, fraudlent behaviour involves gaining control of customer accounts, attempting to transfer funds to another account and then cashing out. |
| isFlaggedFraud | Numeric | Identifies transactions that go over a predetermined limit. Any transaction in this dataset that involves more than 200,000 transfers is marked as possibly fraudulent. |

**Data Preparation**

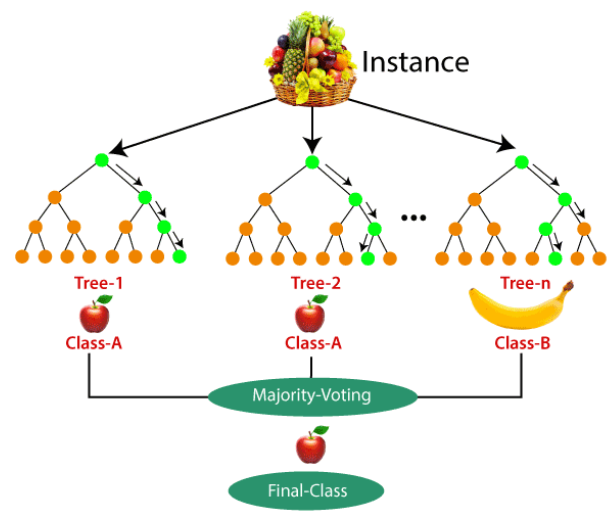
The dataset, we first loaded the data and checked the data frame using pandas looks like. How the data is categorical, numerical values is there in my dataset and checked missing values, which were handled appropriately. In my data there is no missing values. Duplicate rows are removed to ensure the data integrity. This step is crucial as it provided additional insights that could help in identifying fraudulent transactions. I have applied Label Encoding for three features like type, nameOrig, nameDest. label Encoding use for converting into categorical feature to numerical features. It can understand machine easily. After the remaining features like amount, oldbalanceOrg, newbalanceOrig, oldbalanceDest, and newbalanceDest were standardized to sure that every feature made an equal contribution to the model, with a mean of 0 and a standard deviation of 1. After then, the dataset was divided into training and testing sets, with 20% going towards testing and 70% going towards training. when I have applied SMOTE to oversample of the minority class. When I have applied SMOTE the implementing the Machine learning models. In my dataset the fraud is out of 6lakh 50 rows 8213 is Fraud. Finally, the prepared data was ready to model training and evaluation.

**Modelling**

The modelling phase, I focus on training and evaluating machine learning and deep learning model to detect fraudulent transactions. The steps involved:

I start by selecting appropriate machine learning algorithms like Logistic Regression, Decision Tree, Random Forest, Naïve Bayes, XGBoostclassifier. In Deep learning models like Artificial Neural Network (ANN), Conventional Neural Network (CNN), Long short-term memory (LSTM).

**Random Forest**

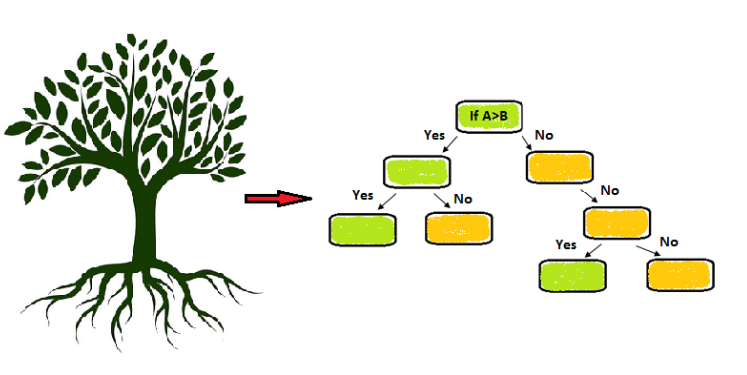


# Figure- 5: Random Forest image

Random Forest algorithm

In order to create a class that is the model of the individual trees classification and regression classes, the ensemble learning techniques known as ‘random forest’ generates many decision trees during training (Mondal.A ,Sarkar.S etc 2024). To increase accuracy and resilience, it integrates the predictions of several models. A decision tree is represented by each tree in the forest. This type of model creates a structure resembling a decision tree by segmenting the data into subsets based on future values (Chen, X., Mao, Z. etc 2024).

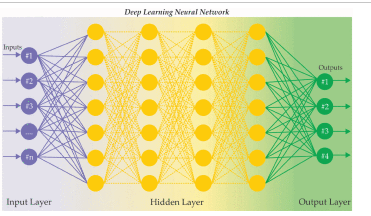
**Decision Tree**



# Figure-6: Decision Tree image

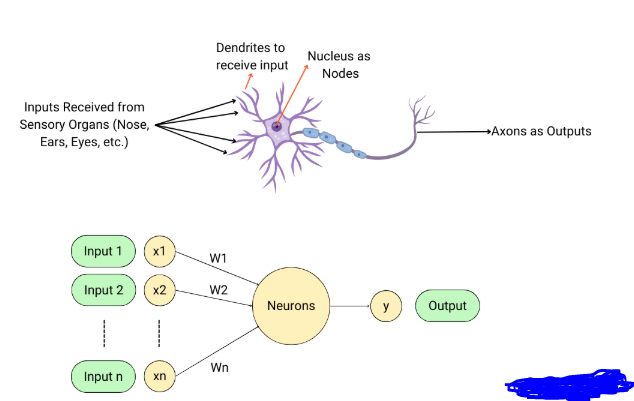
A common supervised technique for classification and regression applications is the decision tree. When the data is separated into smaller groups, a decision tree-like model is created according to the value of the input attributes ( Khemissi.L , Hajjaji.A etc 2024). The decision tree’s top node, which stands for the entire the dataset. It is divided into a minimum of two uniform groups. An attribute test is represented by each decision node.

**Artificial Neural Network (ANN)**



# Figure-7: Artificial Neural Network image

A computational model known as an artificial neural network (ANN) was developed with inspiration from the way information is processed by biological neural networks in the human brain (Wang, G., Bao, H. 2024). Applications for artificial neural networks (ANNs) are numerous and include pattern recognition, regression, and classification. The fundamental components of an ANN accept data, process it,

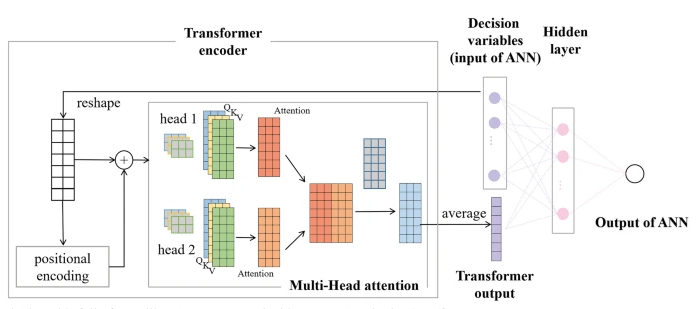


# Figure-8: Artificial Neural Network inside the image

Inside of ANN node within a hidden layer

Each neuron receives inputs from the previous layer. These inputs can be the raw data (for the first hidden layer) or the outputs of neurons from the previous hidden layer ( Bhambu.A , Suganthan .P etc 2024). The weights assigned to each input establish its relative relevance. The method of training involves learning how to use weights.

Genetic Algorithm with Neural Network (GANN)



# Figure-9: Genetic algorithm with Neural Network (GANN)

Natural selection serves as the inspiration for optimisation strategies known as genetic algorithms. They evolve solutions to optimisation issues through methods like crossover, mutation, and selection over generations (Dang, D.C., Opris, A.etc 2024). Each neural network is represented as a chromosome. This can include the networks architecture and the strength of the neural connections.

Each neural network is trained on the dataset, and its performance is evaluated using a fitness function. This fitness function might be based on accuracy, loss or any other relevant metric (Joshi, A.A. and Aziz, R.M., 2024). Neural networks with the highest performance are chosen to be parents to the following generation. Selection techniques like include tournament selection and roulette wheel selection, etc.

The new generation of neural networks is evaluated, and the process repeats for a number of generations or until a satisfactory solution is found.

**Evaluation Metrics**

Predicting categorical labels for input data is a requirement of classification, and it is essential to have strong assessment metrics that can be capture the subtleties of these predictions. These metrics allows to assess the quality of our models by assessing their predictive abilities in different classes and quantifying their performance in terms of accuracy, precision, recall score, f1\_score and confusion matrix etc.

**Accuracy**

The important evaluation parameter that sheds light on how well categorisation models are performing is accuracy. It indicates how accurate a model’s predictions are generally. By comparing the number of correctly instanced instances to the total number of instances in any classification dataset, accuracy is ascertained. The mathematical formula of accuracy is:

**Number of correct predictions**

**Accuracy= -------------------------------------------**

**Total number of instances**

**Precision**

One of the most important metrics for evaluating a classification model’s performance is precision. It calculates the percentage of the models total positive predictions that are true positive predictions. The precision of a model is its ability to minimise false positive predictions while properly identifying positive cases. The precision in mathematics formula is:

**True positives**

**Precision = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**(True positives + False positives)**

**Recall**

A key evaluation parameter for evaluating a classification models’ performance is recall, which is also referred to as sensitivity or true positive rate. It calculates the percentage of genuine positive predictions among all true positive cases.

The Mathematical formula of recall is:

**True positive**

**Recall = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**(True positive + False positive)**

**F1\_score**

A popular assessment tool that provides a fair evaluation of a classification model’s performance by combining precision and recall into one metric is the F1\_score. The mathematical formula of f1 score is:

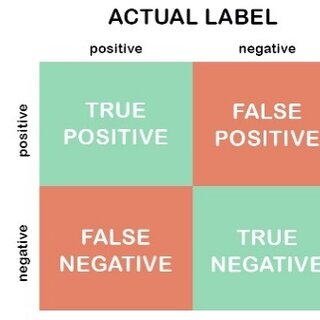
**Precision \* Recall**

**F1\_score = 2 \* \_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**(Precision + Recall)**

**Confusion Matrix**

A useful tool for classification problems is the confusion matrix, which offers a structured and comprehensive overview of the models’ predictions and the actual class labels of dataset instances. The number of accurate and inaccurate predictions is graphically shown, which aids in assessing a classification models performance.

**Confusion matrix:** 

# Figure-10: Confusion matrix A close up of numbers Description automatically generated

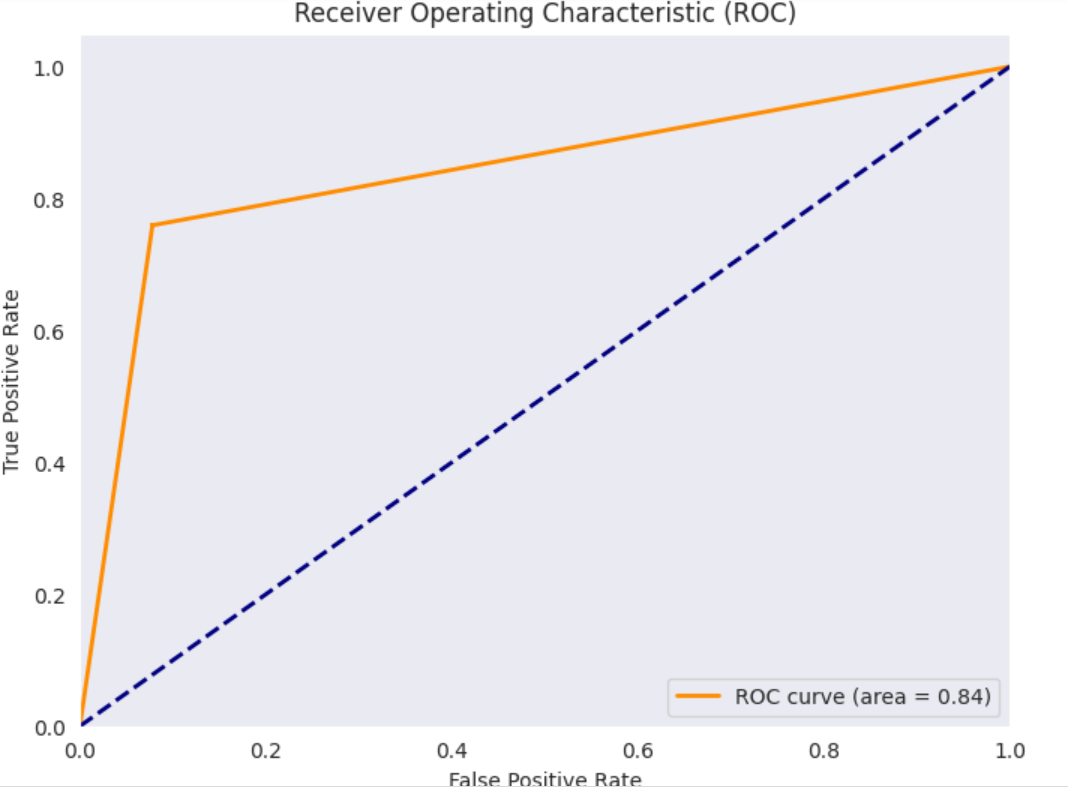
In the above figure is True positive is 3832783. It is False positive is 233692. In the confusion matrix False positive is 161153. It is True Negative is 3905976.

**AUC & ROC**

Common evaluation metrics for binary classification tasks include the receiver operating characteristics (ROC) curve and the area under the curve (AUC). They shed light on a classification models effectiveness and capacity for discrimination.

A visual depiction of the model’s performance at various categorisation thresholds is the ROC curve. For various threshold setting, the true positive rate (TPR) and false positive rate (FPR) are plotted against one another.

The image illustrates an example of desired AUC and ROC



# Figure-11: ROC curve image in Logistic Regression

A model with the highest AUC score performs better and has a higher capacity for class discrimination. A greater positive rate (TPR) compared to false positive rate (FPR) across different classification criteria is shown by the ROC curve when it is closer to the top left corner. As a result, the model is better able to balance minimising false positive mistakes with accurately identifying positive events.

**Deployment**

Enhancing cybersecurity for financial transactions using AI, ML and DL involves several important steps. To begin with, identify common threats like fraud and spam. For instance, AI can detect usual transaction patterns indicative of fraud. Collect and pre-process transaction data, ensuring it’s cleaning and relevant. Train models using algorithms. I have explained in the above modelling and validating from hardness. For instance, a neural network can learn to identify fraudulent transactions by identifying historical data. Deploy these models in real-time monitoring and anomaly detection, integrating them securely with existing systems.

Continuously monitor and update the models with the new data. Educate users on safe practices. Such as identifying spam attempts. Ensure compliance with regulations like General data production regulations. This approach aims to significantly decrease fraud and improve the security of financial transactions, providing a safer environment for users and financial intuitions.

Diagram of a diagram of a model

Description automatically generated

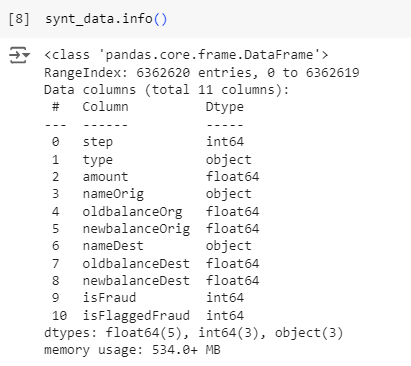
# Figure-11: pre-processing of the data

# 4Experiments

In this section, every experiment carried out for this study has documentation. The proposed model, which shows each step of the process in section 3.3 overview of the process model, serves as the basis for the tests that were conducted. Python is the programming language utilised in this study and google colab is the integrated development environment. This research used the synthetic-data dataset applied to the datasets which create the financial fraud detection classification model where I implemented in the in code. I discussed below.

## 4.1 Data understanding

To begin with, the synthetic dataset to understand the datasets like the data types of attributes, the crucial of the variables, and the feature selections of the attributes. The datasets are subjected to the following data preparation methods.



# Figure-12: Information of dataset

In the above synthetic dataset information

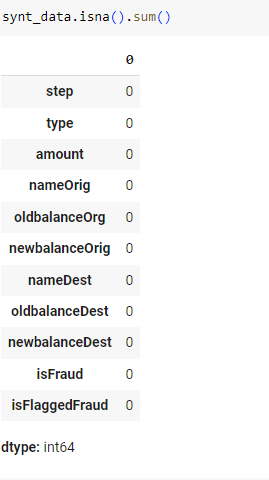
The information about the variables in the datasets is displayed in the above figure. In this dataset contains 10 attributes. The dataset contains combinations of categorical and numerical values.

## 4.2 Data preparation

### 4.2.1 Data Pre-processing

In any project involving data analysis or machine learning, data preparation is an essential stage. It entails converting the unprocessed data into an organised and functional format.

In the below the few key steps involved in data preparation.



# Figure:13: image illustrates the dataset there is no missing values

In the above figure, in my dataset there is no missing values. The data information is clean and good

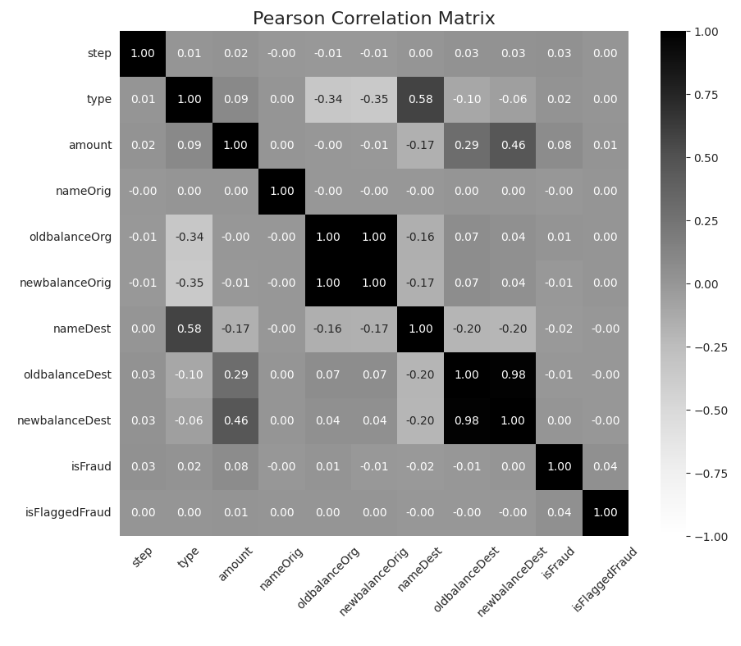
## 4.2.2 Data Visualisation

Data visualisation is the process of representing information and data graphically. Data visualisation tools offer an easily comprehensible means of viewing and comprehending data, trends, outliers, and patterns through the use of visual elements such as charts, graphs and maps.

To begin with, benefits of data visualisation is improved understanding. Visuals make it easier to grasp difficult concepts or identify new patterns. For instance, a line graph represents can show trends over time while a bar chart can compare quantities across different categories.

In addition, Data visualizations also aids in better decision-making. By clearly and summary presenting the data, it allows stake holders to make decisions quickly.

There are many visualizations tools available for data visualizations such as tableau, power BI and python libraries like matplotlib and seaborn. Every tool has advantages and the best one to choose will depend on the particular requirements of the project.



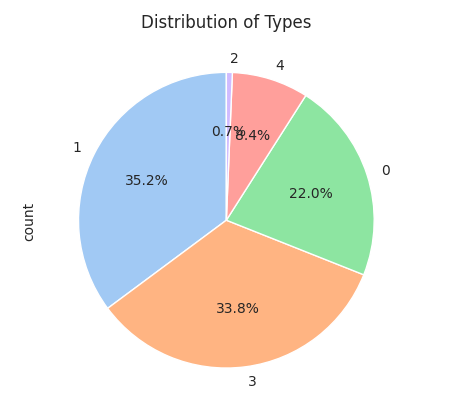
# **Figure:14 Correlation Plot**

A square table with rows and columns representing distinct variables is called a correlation matrix. The correlation coefficient between the variables in the respective row and column is represented by the value in each table cell. The correlation coefficient ranges from -1 to +1:

* +1 indicates a perfect positive correlation
* -1 indicates a perfect negative correlation
* 0 indicates no correlation

It helps in identifying the strength and direction of relationships between variables. Summaries of large dataset and highlight patterns.

In the below figure, a pie chart is a statical shows that is round and has been divided into slices, represents data in a round format. Each slice of a pie corresponds to a category’s proportion of a total. The entire pie chart represents 100% of the data, each slice size is proportional to its percentage of the total.



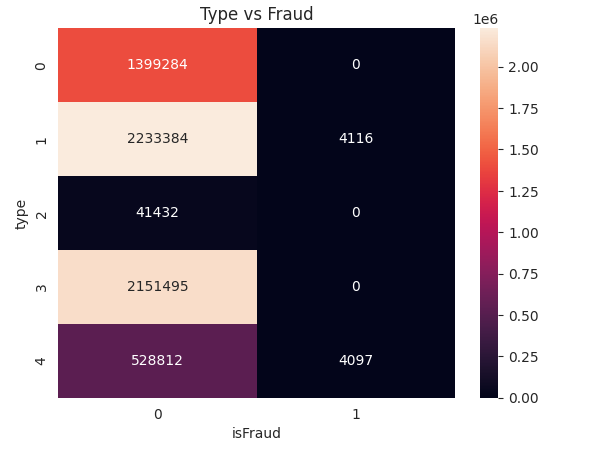
# Figure:15: Pie chart

In the above figure, the pie chart shows each slices represents a category’s contribution to the total.

0 means- Cash In-22% 1 means-Cash Out-35.2%, 2 means-Debit-0.7%, 3 means- Payment-33.8%, 4 means-Transfer-8.4%

Easy to understand the data at a glance, making them great for presentation of reports.

A heatmap is a visual representation of data in which the colours correspond to the individual values within a matrix. It gives a quick visual overview of the data, which facilities the identification of trends, correlations and outliers.



# Figure:16: Heatmap of the data

In the above figure shows the types of fraud and non-fraud in our financial datasets. Cash-Out fraud transactions involves unauthorized transactions where fraudsters use stolen credit card details, bank account information, or other financial credentials to withdraw money. This can occur through various methods, including ATM withdrawals, online transactions, or point-of-sale-purchases.

In the above figure shows Cash-out transactions happened 4116 transactions as a fraud. Another time while doing the transfer to one to another person. Incorrect account number or IFSC code. It went unknown person 4097 transactions is fraud.

The bar chart is a graphical representation of data using rectangular bars. The length of the height of each bar is proportion to the value it represents.

A graph of different types of blue bars

Description automatically generated

# Figure:17: Count of different types of transactions

In the above shows the bar chart appears to represent the frequency distribution of the types of labels different categories. 0 is represents the CASH IN. Another bar chart shows the 1 represents the CASHOUT type in the synthetic dataset. Followed by 2 is represents the DEBIT transactions in my dataset. Followed by 3 represents the PAYMENT type in the dataset. Followed by TRANSFER transactions in my dataset.

A pie chart with numbers and a red circle with Crust in the background

Description automatically generated

# Figure:18: Pie chart shows each type of transactions

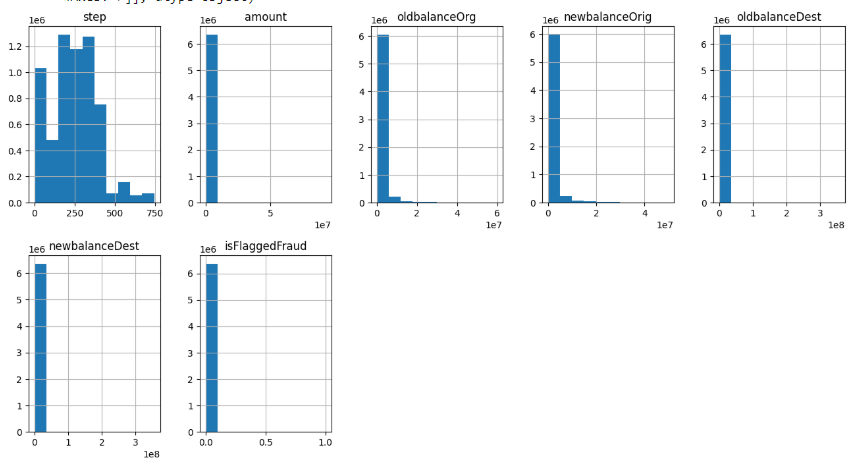
In the above image, the pie chart with four different colour sections, each labelled with a percentage. The largest is red and accounts for 35.2%. It is CASH OUT in my dataset. Followed by another section of the chart is yellow section at 33.8%, it indicates the PAYMENT. In another section orange. It indicates the 22.0%. this part is CASHIN. Another section is green colour is 8.4% of the dataset. It indicates the TRANSFER. Final section is tiny part of the pie chart is 0.7% .it indicates the DEBIT. In the above shows types of transactions in my dataset.

A screenshot of a graph

Description automatically generated

# Figure:19: Target column is Fraud and Non-Fraud

In the above bar chart shows the is Fraud and non-fraud fraud transactions in my dataset. The target feature is is Fraud in my dataset.



# Figure:20: Histograms study each feature in my dataset

A histogram is a visual depiction of how numerical data is distributed. It is made by counting the number of data points that fall into each bin after the data range is divided into intervals called bins.

Every feature of each column in the dataset is depicted in the above figure. !! feature of the bins in the above graphic are included in the data set.

## 4.2.3 Data Transformation

Converting data from one format or structure to another is known as data transformation. In order to ensure that data is clean, consistent and suitable for a range of applications, this phase is essential to data management and analysis.

To begin with, Modifying the format, structure or values of data is known as data transformation. This process is essential for integrating data from different sources, preparing for data analysis and improving data quality. In my dataset, I have applied 4 features are categorical. I have applied for Label Encoding. I have converted into categorical to numerical values. It can machine understand the data easily. It is main crucial step for data transformation.

In my dataset, When I have applied SMOTE (synthetic-minority oversampling technique). In my dataset has a significant imbalance between classes (e.g.: many more legitimate transactions than fraudulent ones), More often, the model might start to favour the majority class, which would result in high accuracy but subpar performance on the minority class.

When deep learning models are trained on unbalanced data, they are more likely to predict the majority class, ignoring the minority class. This result in low recall and precision for the minority class, which is often the class of interest in applications like fraud detection.

In addition, in this dataset are splitting into 80% training and 20% testing is a common practice to evaluate the performance of model. I have applied Power transformer for Logistic regression. The power transformer is used to make data more Gaussian-like which can be beneficial for many machines learning model’s algorithms. It helps stabilize variance and minimize skewness and making the data more suitable for modelling. Yeo-Johnson transformation works with both positive and negative data, making it more flexible.

Furthermore, by eliminating the mean and scaling to unit variance, the standard scaler standardises characteristics. It is a crucial step for deep learning models because it ensures each features contribute equally to the models learning process. Standardizing the data helps in faster mixing during training and can lead to improved model performance, particularly for algorithms like neural networks and other algorithms that are sensitive to feature sizes.

In my dataset, the financial transactions are isfraud is 8213 is equal to 0.13% and good transactions are 99.87% out of 100%.

A blue circle with black numbers

Description automatically generated

# Figure:21: Pie chart of the fraud transactions and Non fraud transactions in the data

The image illustrates a pie chart with two segments. The larger segment is blue, covering almost the entire chart and is labelled ’99.87%’ and ‘0’. The smaller segment is orange, representing a small fraction of the chart, labelled’0.13% and ‘1’.

The distribution of classes in a dataset, possibly in the context of fraud detection where one class (is Fraud) is much less frequent than the other (nonfraud).

When I have applied SMOTE (synthetic minority oversampling technique)

A blue and orange circle with numbers

Description automatically generated

# Figure:22: When I have applied SMOTE

The figure illustrates a pie chart divided into two equal types, one coloured orange and the other colour is blue, each part is representing 50% of the whole. When I have applied SMOTE (synthetic minority oversampling technique). The blue colour part ‘0’. It indicates Fraud. Another colour part is representing the ‘1’. It indicates the non-fraud. I did balance the data apply SMOTE. We can apply such as under sampling, oversampling and ADASYN.

## 4.3 Model training

In machine learning, model training is a basic procedure where an algorithm learns from data to anticipate outcomes. To ensure accurate and dependable performance in real-world applications, the objectives is to develop a model that can generalise well to new, unseen data.

This involves cleaning and transforming the data to ensure it is suitable for training. Handling missing values, standardising or normalising features, and dividing the data into training and testing sets are typical procedures.

The synthetic dataset designed to simulate real-world financial transactions data, serves as the foundation for training various models.

Data pre-processing steps, including data cleaning, feature engineering and splitting the dataset into training, validation and test sets, ensure the quality and reliability of the data.

Multiple models, such as logistic regression, decision trees and neural networks are selected based on their suitability for detecting anomalies in financial transactions.

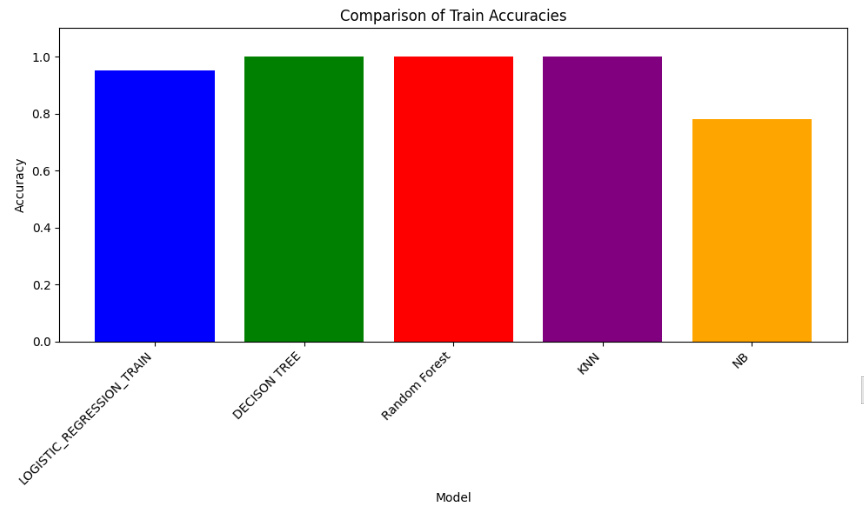
The training process involves hyperparameter tuning for deep learning models such as Artificial Neural Network, Conventional neural Network, long short-term memory for tuning using methods like grid search and random search.

The application of training algorithms such as gradient descent and backpropagation. Evaluation metrics, including accuracy, precision, recall and F1-score are used to assess the model performance.

|  |  |
| --- | --- |
| **Model** | **Hyperparameter optimisation** |
| Random  Forest |  |
| LSTM |  |

In the above Random Forest classifier

## 4.4 Model Evaluation

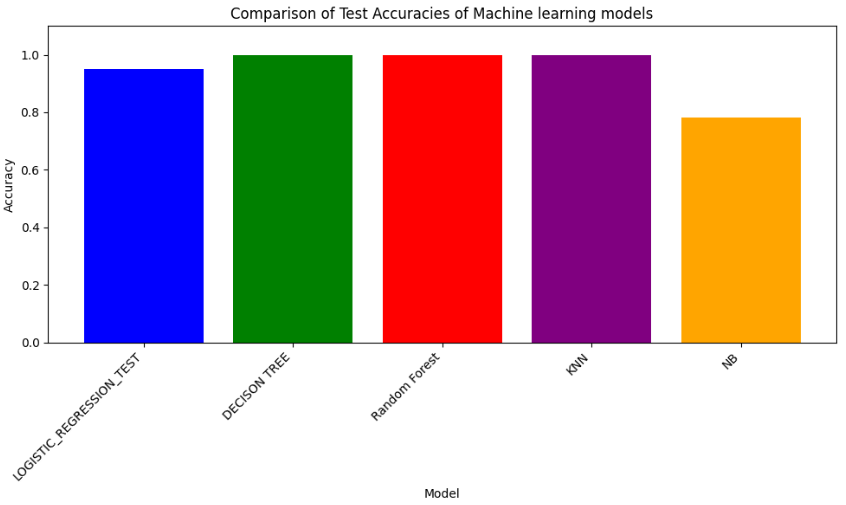


# Figure:23 Accuracy of Machine learning Models Training

In the above figure shows the machine learning models such that Logistic regression of train accuracy is 95.158%. In normalization technique such that power transformer instead of standard scaler in data normalization. The data balancing the technique is SMOTE for machine learning models.

The decision tree of accuracy is 100%. In data normalization technique is Standard scaler. the data balancing. I have synthetic minority oversampling technique. The target feature is 50% Fraud and 50% non-fraud of this dataset. The RandomForestclassifier algorithm accuracy is 100%. The data normalization technique is Standard scaler. The data balancing technique is Synthetic minority oversampling technique.

The K-nearest neighbour train accuracy is 100%. The data normalization technique is standard scaler. The data balancing the data is synthetic minority oversampling technique of KNN algorithm. The Naïve bayes algorithm train accuracy is 78%. The data normalization technique is standard scaler is rescaling the features such that they have the properties of a standard normal distribution with a mean of a zero and a stand deviation of one in standard scaler. The data balancing is synthetic minority oversampling technique. Out of 5 machine learning algorithms Decision tree classifier, Random forest classifier and K-nearest neighbour.



# Figure:24: Test accuracy of Machine learning models

The testing accuracy of Machine learning models such that logistic regression is accuracy is 95.157%. The data normalization technique like power transformer like box-cox aim to make the make data as normal as possible. In my dataset, the balancing the data I have applied SMOTE (synthetic minority oversampling technique). In decision tree classifier the accuracy is 99.94%.

The data normalization technique is standard scaler of normal distribution with a mean of zero and a standard deviation of one. The balancing data, I have applied synthetic minority oversampling technique and their minority class and their neighbours. The Random Forest classifier algorithm test accuracy is 99.95%. The data normalization, I have applied Standard scaler of this algorithm. The data balancing, I have applied synthetic minority oversampling technique of this algorithm.

The K-nearest neighbour classification algorithm accuracy test is 100%. The data normalization, I have applied standard scaler for data normalization of data. The balancing the data, I have applied synthetic minority oversampling technique.

The algorithm of Naïve bayes is accuracy is 78%. In data normalization, I have applied Standard scaler. The balancing data, I have applied synthetic minority oversampling technique.

**Train Accuracy of Machine learning models**

# Table:3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Accuracy** | **Precision** | **Recall** | **F1-score** | **AUC score** |
| Logistic regression + Power Transformer +SMOTE | 95.13% | 94% | 96% | 95% | 95.158% |
| Decision Tree classifier + Standard Scaler+ SMOTE | 100% | 100% | 100% | 100% | 100% |
| Random Forest classifier + Standard Scaler + SMOTE | 100% | 100% | 100% | 100% | 100% |
| KNN+ Standard scaler + SMOTE | 100% | 100% | 100% | 100% | 100% |
| Naïve Bayes + Standard Scaler + SMOTE | 78.06% | 77.08% | 79.84% | 78.44% | 78.06% |

**Test Accuracy and Evaluation metrics of Machine learning models**

# Table:4

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Models** | **Accuracy** | **Precision** | **Recall Score** | **F1-score** | **AUC**  **score** |
| Logistic Regression + Standard scaler + SMOTE | 95.14% | 94.13% | 96% | 95.19% | 95.158% |
| Decision Tree classifier + Standard Scaler + SMOTE | 99.9% | 99.9% | 99.96% | 99.94% | 99.944% |
| Random Forest classifier + Standard scaler + SMOTE | 99.95% | 99.91% | 99.99% | 99.95% | 99.95% |
| KNN + Standard scaler + SMOTE | 99.95% | 99.91% | 99.99% | 99.95% | 99.957% |
| Naïve Bayes + Standard scaler + SMOTE | 78.06% | 77.08% | 79.84% | 78.44% | 78.060% |

**Training and Test accuracy of Deep learning models**

# Table:5

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Test accuracy** | **Train accuracy** | **Hyperparameter tuning** |
| Artificial Neural Network (ANN) + Standard scaler | 99.964% | 0.9996342062950134 | 0.9996424317359924 |
| Conventional Neural Network + Standard scaler | 99.968% | 0.9996681809425354 | 0.9996880292892456 |
| Long short-term memory (LSTM) +Standard scaler | 99.963% | 0.9996257424354553 | 0.9996314644813538 |

# 5.0 Conclusion and future works

## 5.1 Introduction

In today’s digital age, the financial sector is increasingly reliant on technology to facilitate transactions and mange sensitive data. However, this reliance also exposes financial institutions to growing array of cyber threats. Cybercriminals are continually evolving their tactics, making it imperative for the financial industry to adopt advanced cybersecurity measures.

Artificial intelligence (AI), Machine learning (ML) and Deep learning (DL) have emerged as powerful tools in the fight against cybercrime. These technologies offer the ability to analyse vast amounts of data, identify patterns and predict potential threats with unprecedented accuracy. By leveraging AI, ML and DL financial institutions can enhance their cybersecurity defences, detect anomalies in real-time and respond to threats more effectively.

This report explores the integration of AI, ML and DL in enhancing cybersecurity for financial transactions. It examines the current landscape of cyber threats, the role of these technologies in mitigating risks, and the challenges associated with their implementation. Through a comprehensive analysis, this report aims to provide insights how AI, ML and DL can be harnessed to protect financial transactions and ensure the integrity and security of financial system.

## 5.2 Objective Achievement

By implementing AI and ML algorithms, the project successfully developed models capable of detecting and preventing various cyber threats in real-time. These models were trained on extensive datasets, enabling them to identify anomalies and potential security breaches with high accuracy. The Fraud detection integration of DL techniques allowed for the creation of sophisticated fraud detection systems. These systems were able to analyse transaction patterns and user behaviour, identifying fraudulent activities with minimal false positives.

The automated response system achieved the development of automated response systems that utilize AI to respond to detect threats swiftly. These systems can isolate compromised accounts, alert relevant authorities, and initiate countermeasures to mitigate the impact of cyber-attacks. Through the application of advanced encryption algorithms and secure data handling practices, the project ensured that sensitive financial data remains protected from unauthorized access and breaches.

The AI, ML and DL models developed during the project were designed to be scalable and adaptable. This ensures that the cybersecurity measures can evolve alongside emerging threats and changing technological landscapes. The project also focused on educating users about cybersecurity best practices. By raising awareness and providing training, the project contributed to a more security-conscious user base, future enhancing the overall security of financial transactions.

## 5.3 Limitations and Future works

The effectiveness of AI, ML and Dl models heavily depends on the quality and quantity of data available. Inadequate or biased data can lead to inaccurate predictions and missed threats. Ensuring access to comprehensive and high-quality datasets remain a challenge. In computational resources implanting advanced Ai and DL models require significant computational power and resources. Smaller financial institutions may face difficulties in acquiring the necessary infrastructure, limiting their ability to fully leverage these technologies.

While AI and DL models can achieve high accuracy, their complexity often makes them difficult to interpret. This lack of transparency can hinder truest and understanding mong stakeholders, particularly in regulatory environments where explain ability is crucial. The evolving threat landscape in cyber threats are constantly evolving and attackers are becoming more sophisticated. AI and ML models need continuous update and retraining to stay effective, which can be resource-intensive and challenging to maintain.

The use of AI in cybersecurity raises various regulatory and ethical issues, including data privacy, consent and potential biases in decision making. Navigating these concerns requires careful consideration and adherence to legal and ethical standards.

What I am discussed in the research questions related to the research gap. Here the explanation:

The financial sector is a prime target for cybercriminals due to the vast amounts of sensitive data and financial assets it handles. In the rapidly evolving digital landscape, the financial sector faces an increasing array of cyber threats, necessitating robust and adaptive cybersecurity measure. Traditional security methods, while foundational, often fall short in addressing sophisticated and emerging threats such as phishing, ransomware, data breaches, DDoS attacks, advanced persistent threats (APTs) and supply chain attacks.

These conventional approaches are typically reactive, struggle with scalability and rely heavily on human intervention, which can introduce errors and inconsistences. To counter these challenges, financial institutions are turning to advanced technologies like Artificial intelligence (AI), Machine learning (ML) and Deep learning (DL).

These technologies offer proactive and intelligent solutions by analysing vast datasets, identifying patterns and predicting potential threats with high accuracy. AI and ML models enhance threat detection and prevention, while DL techniques bolster fraud detection by analysing transaction patterns and user behaviour. Automated response systems, powered by Ai, enable swift reactions to detected threats, isolating compromised accounts and initiating countermeasures.

Despite their promise, these technologies face limitations, including data quality and availability, computational resource demands, model interpretability and evolving threat landscapes. Future work should focus on improving data collection, developing explainable AI models, creating scalable and cost-effective solutions, and establishing robust ethical and regulatory frameworks.

By leveraging AI, ML and Dl financial institutions can significantly enhance their cybersecurity defences, protect sensitive data, and ensure the integrity and security of financial transactions in an increasingly digital world.

The following answer is related to the research question 2:

Integrating AI, ML and DL technologies into existing cybersecurity frameworks within financial institutions involves several strategic steps to enhance threat detection, prevention and response capabilities. Financial institutions must start by collecting and pre-processing large volumes of data from various sources, such as transaction records and network traffic, to train AI and ML models. This model can then be deployed to monitor and analyse data in real-time, identifying anomalies and potential threats with high accuracy. DL techniques can be used to develop sophisticated fraud detection systems that analyse transaction patterns and user behaviour significantly reducing false positives.

Ai-powered automated response systems can swiftly react to detect threats, isolating compromised accounts and initiating countermeasures. Additionally, AI and ML can enhance existing security tools like firewall and intrusion detection systems by adding an extra layer of intelligence and automation. Continuous learning mechanisms ensure that AI models adopt to new and emerging threats, maintaining the effectiveness of the cybersecurity framework.

Collaboration across various departments within in the financial institutions including IT, compliance and risk management is crucial to align Ai and ML models with regulatory requirements and organizational policies. By following these steps, financial institutions can significantly enhance their cybersecurity defences, protect sensitive data, and ensure the integrity and security of financial transactions in an increasingly digital world.

The following answer is related to the research question no 3:

Implementing AI, ML and DL technologies for cybersecurity in financial institutions present several challenges. One major challenge is the quality and availability of data, as inconsistent or biased data can lead to inaccurate predictions.

Additionally, the significant computational resources required for advanced AI and DL models can be a barrier, especially for smaller institutions. The complexity and lack of interpretability of these models can also hinder trust and understanding among stakeholders, particularly in regulatory environments. The constantly evolving threat landscape necessitates continuous updates and retraining of AI models, which can be resource-intensive. Regulatory and ethical concerns, such as data privacy and potential biases, further complicate the implementation process. Integrating AI technologies with legacy systems and addressing skill shortages in AI and cybersecurity expertise are additional hurdles.

To mitigate these challenges, financial institutions can improve data quality through rigorous validation processes, leverage cloud-based solutions for scalable resources and focus on developing explainable AI models. Continuous learning mechanisms and robust ethical frameworks can help adapt to new threats and ensure responsible use of AI.

Enhancing integration capabilities and investing in upskilling programs can also facilitate smoother implementation and better utilization of AI, ML and DL technologies in cybersecurity frameworks.

# References

1. Roshanaei, M., Khan, M.R. and Sylvester, N.N., 2024. Navigating AI cybersecurity: evolving landscape and challenges. *Journal of Intelligent Learning Systems and Applications*, *16*(3), pp.155-174.
2. Hilal, W., Gadsden, S.A. and Yawney, J., 2022. Financial fraud: a review of anomaly detection techniques and recent advances. *Expert systems With applications*, *193*, p.116429.
3. Ranjan, P. and Dahiya, S., 2021. Advanced threat detection in api security: Leveraging machine learning algorithms. *International Journal of Communication Networks and Information Security*, *13*(1).
4. Wang, S., Asif, M., Shahzad, M.F. and Ashfaq, M., 2024. Data privacy and cybersecurity challenges in the digital transformation of the banking sector. *Computers & security*, *147*, p.104051.
5. Chatterjee, P., Das, D. and Rawat, D.B., 2024. Digital twin for credit card fraud detection: Opportunities, challenges, and fraud detection advancements. *Future Generation Computer Systems*
6. Wang, S., Asif, M., Shahzad, M.F. and Ashfaq, M., 2024. Data privacy and cybersecurity challenges in the digital transformation of the banking sector. *Computers & security*, *147*, p.104051.
7. Rana, M.U., Shah, M.A., Alnaeem, M.A. and Maple, C., 2024. Ransomware Attacks in Cyber-Physical Systems: Countermeasure of Attack Vectors Through Automated Web Defenses. *IEEE Access*.
8. Abdelkader, S., Amissah, J., Kinga, S., Mugerwa, G., Emmanuel, E., Mansour, D.E.A., Bajaj, M., Blazek, V. and Prokop, L., 2024. Securing modern power systems: Implementing comprehensive strategies to enhance resilience and reliability against cyber-attacks. *Results in engineering*, p.102647
9. Mahboubi, A., Luong, K., Aboutorab, H., Bui, H.T., Jarrad, G., Bahutair, M., Camtepe, S., Pogrebna, G., Ahmed, E., Barry, B. and Gately, H., 2024. Evolving techniques in cyber threat hunting: A systematic review. *Journal of Network and Computer Applications*, p.104004.
10. Morić, Z., Dakic, V., Djekic, D. and Regvart, D., 2024. Protection of Personal Data in the Context of E-Commerce. *Journal of cybersecurity and privacy*, *4*(3), pp.731-761.
11. Sarker, I.H., Janicke, H., Mohsin, A., Gill, A. and Maglaras, L., 2024. Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects. *ICT Express*.
12. Kavitha, D. and Thejas, S., 2024. AI Enabled Threat Detection: Leveraging Artificial Intelligence for Advanced Security and Cyber Threat Mitigation. *IEEE Access*.
13. Gulyas, O. and Kiss, G., 2023. Impact of cyber-attacks on the financial institutions. *Procedia Computer Science*, *219*, pp.84-90.
14. Mallick, M.A.I. and Nath, R., 2024. Navigating the Cyber security Landscape: A Comprehensive Review of Cyber-Attacks, Emerging Trends, and Recent Developments. *World Scientific News*, *190*(1), pp.1-69.
15. Ding, W., Abdel-Basset, M., Hawash, H. and Ali, A.M., 2022. Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Information Sciences*, *615*, pp.238-292.
16. Hilal, W., Gadsden, S.A. and Yawney, J., 2022. Financial fraud: a review of anomaly detection techniques and recent advances. *Expert systems With applications*, *193*, p.116429.
17. Gandhar, A., Gupta, K., Pandey, A.K. and Raj, D., 2024. Fraud Detection Using Machine Learning and Deep Learning. *SN Computer Science*, *5*(5), pp.1-10.
18. Aderamo, A.T., Olisakwe, H.C., Adebayo, Y.A. and Esiri, A.E., 2024. AI-driven HSE management systems for risk mitigation in the oil and gas industry. *Comprehensive Research and Reviews in Engineering and Technology*, *2*(1), pp.1-22.
19. Bueno, L.A., Sigahi, T.F., Rampasso, I.S., Leal Filho, W. and Anholon, R., 2024. Impacts of digitization on operational efficiency in the banking sector: Thematic analysis and research agenda proposal. *International Journal of Information Management Data Insights*, *4*(1), p.100230.
20. Umar, M.A. and Lano, K., 2024. Advances in automated support for requirements engineering: a systematic literature review. *Requirements Engineering*, pp.1-31.
21. Sivamayilvelan, K., Rajasekar, E., Balachandran, S., Kotecha, K. and Vairavasundaram, S., 2024. Data-driven approach for identifying the factors related to debt collector performance. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(4), p.100385.
22. Nkambule, D.S., Twala, B. and Pretorius, J.H.C., 2024. Effective Machine Learning Techniques for Dealing with Poor Credit Data. *Risks*, *12*(11), p.172.
23. Muñoz-Cancino, R., Bravo, C., Ríos, S.A. and Graña, M., 2023. On the dynamics of credit history and social interaction features, and their impact on creditworthiness assessment performance. *Expert Systems with Applications*, *218*, p.119599
24. Zhu, Q., Zhou, X. and Liu, S., 2023. High return and low risk: Shaping composite financial investment decision in the new energy stock market. *Energy Economics*, *122*, p.106683.
25. Gupta, A. and Agarwal, P., 2024, May. Enhancing sales forecasting accuracy through integrated enterprise resource planning and customer relationship management using artificial intelligence. In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-6). IEEE.
26. Karthikeyan, T., Govindarajan, M. and Vijayakumar, V., 2023. An effective fraud detection using competitive swarm optimization based deep neural network. *Measurement: Sensors*, *27*, p.100793.
27. Mienye, E., Jere, N., Obaido, G., Mienye, I.D. and Aruleba, K., 2024. Deep Learning in Finance: A survey of Applications and techniques. *AI*, *5*(4), pp.2066-2091.
28. Gambacorta, L., Huang, Y., Qiu, H. and Wang, J., 2024. How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. *Journal of Financial Stability*, p.101284.
29. Parente, M., Rizzuti, L. and Trerotola, M., 2024. A profitable trading algorithm for cryptocurrencies using a Neural Network model. *Expert Systems with Applications*, *238*, p.121806.
30. Bhatore, S., Mohan, L. and Reddy, Y.R., 2020. Machine learning techniques for credit risk evaluation: a systematic literature review. *Journal of Banking and Financial Technology*, *4*(1), pp.111-138
31. Shah, V., 2021. Machine learning algorithms for cybersecurity: Detecting and preventing threats. *Revista Espanola de Documentacion Cientifica*, *15*(4), pp.42-66.
32. Mishra, S., 2023. Exploring the impact of AI-based cyber security financial sector management. *Applied Sciences*, *13*(10), p.5875.
33. Alshaikh, O., Parkinson, S. and Khan, S., 2024. Exploring perceptions of decision-makers and specialists in defensive machine learning cybersecurity applications: The need for a standardised approach. *Computers & Security*, *139*, p.103694.
34. Geetha, R. and Thilagam, T., 2021. A review on the effectiveness of machine learning and deep learning algorithms for cyber security. *Archives of Computational Methods in Engineering*, *28*(4), pp.2861-2879
35. Choudhury, A., Mondal, A. and Sarkar, S., 2024. Searches for the BSM scenarios at the LHC using decision tree-based machine learning algorithms: a comparative study and review of random forest, AdaBoost, XGBoost and LightGBM frameworks. *The European Physical Journal Special Topics*, pp.1-39.
36. Chen, X., Mao, Z. and Wu, C., 2024. Multi-class Financial Distress Prediction Based on Feature Selection and Deep Forest Algorithm. *Computational Economics*, pp.1-40.
37. Khemissi, I., Laajimi, R., Mabrouk, K.B., Hajjaji, A. and Rahmani, M., 2024. An in-depth study by machine learning models of the absorbance properties of stable Ta2O5 nanotubes decorated by PbS nanoparticles for photocatalytic applications. *Materials Research Bulletin*, p.113247
38. Wang, G., Bao, H., Liu, Q., Zhou, T., Wu, S., Huang, T., Yu, Z., Lu, C., Gong, Y., Zhang, Z. and He, S., 2024. Brain-inspired artificial intelligence research: A review. *Science China Technological Sciences*, *67*(8), pp.2282-2296.
39. Bhambu, A., Gao, R. and Suganthan, P.N., 2024. Recurrent ensemble random vector functional link neural network for financial time series forecasting. *Applied Soft Computing*, *161*, p.111759
40. Dang, D.C., Opris, A. and Sudholt, D., 2024. Crossover can guarantee exponential speed-ups in evolutionary multi-objective optimisation. *Artificial Intelligence*, *330*, p.104098.
41. Joshi, A.A. and Aziz, R.M., 2024. A two-phase cuckoo search based approach for gene selection and deep learning classification of cancer disease using gene expression data with a novel fitness function. *Multimedia Tools and Applications*, pp.1-32.

Figures

[A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions | IEEE Journals & Magazine | IEEE Xplore](https://ieeexplore.ieee.org/abstract/document/10403908)