Master Thesis

Master of Science (MSc.)

Major: Data science (DS)

Topic: Detection of credit card fraud using machine learning

Author: Nalla Sumanth

Matriculation Number: 59954796

First supervisor: Prof. Dr Talha Ali Khan

Second supervisor: Mr. Shan Faiz

Submitted on: 28/06/2025

Statutory Declaration

I hereby declare that I have developed and written the enclosed Master Thesis completely by myself and have not used sources or means without declaration in the text. I clearly marked and separately listed all the literature and all the other sources which I employed when producing this academic work, either literally or in coherently. I am aware that violating of this regulation will lead to the failure of the thesis.

Berlin, 28/06/2025 NallaSumanth

Place, Date Signature

Abstract

Credit card fraud is still a key issue in the field of finance where more and more ineffective are old rule-based detection systems facing up against new, adaptive fraudulent methods. The purpose of this study is to create and test machine learning-based framework that can be used to precisely detect fraudulent credit card transactions. Based on an anonymized data evaluated from the Kaggle, the research relies on supervised learning models of Logistic Regression, Random Forest, XGBoost, and Artificial Neural Networks following pre-processing operations, such as SMOTE for the class balancing. Performance is measured using key metrics including precision, recall, F1-score, and AUC-ROC, whereby hyperparameter tuning increases the accuracy and reliability of the models. Findings reveal that simpler classifiers can be overcome by ensemble methods such as XGBoost in the detection of fraud, especially in imbalanced datasets. The research shows that not only does machine learning raise the rate of fraud detection, but it also reduces false positives, which adds both a technical and practical advantage to the financial institutions. Ethical considerations such as privacy of data and fairness of algorithms are also considered. The present study adds to the body of literature in support of AI driven systems of mitigation of financial frauds.

Table of Contents

List of Figures

List of Tables

List of abbreviation

Abbreviation	Full Form
ANN	Artificial Neural Network
AUC	Area Under Curve
CV	Cross-Validation
EDA	Exploratory Data Analysis
F1-Score	Harmonic Mean of Precision and Recall
FN	False Negative
FP	False Positive
KDE	Kernel Density Estimation
LR	Logistic Regression
ML	Machine Learning
PCA	Principal Component Analysis
PR-AUC	Precision-Recall Area Under Curve
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SMOTE	Synthetic Minority Over-sampling Technique
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
XGBoost	eXtreme Gradient Boosting

Chapter 1 Introduction

1.1 Background

1.1.1 Evolution of Payment Systems

Payment systems transform along with societal needs as well as economic requirements through the course of human development. The first type of exchange in early civilizations involved bartering until societies required a standardized payment method. The development of metal coins and eventually paper money made trading operations simpler and easier.

During the twentieth century people gained the ability to send money securely through checks which did not require physical cash. The invention of debit and credit cards brought forward faster payment options which provided consumers with higher ease of use (Tarlin, 2021). Debit cards let users access their own money directly but credit cards enabled borrowing funds up to approved credit limits. New financial innovations restructured money management by enabling faster along with more flexible transaction methods.

Digital payments asserted dominance when the internet became popular during the late 1990s and early 2000s. E-commerce expanded rapidly as credit cards established their position as the preferred online payment method which made credit systems necessary for physical store transactions along with digital commerce (Alzoubi & Ghazal 2022).

1.1.2 Growth of Credit Card Usage

People worldwide have adopted credit cards with rapid expansion as they now function as vital financial instruments. Global credit cards combined with debit cards and prepaid cards exceed 22 billion in total numbers while credit cards represent a major proportion of these transactions (Beju & Făt, 2023). Credit cards have reached widespread adoption in the United States because 82% of adults possess at least one such card.

The advantages that credit cards provide to consumers include the ability to make purchases easily together with flexible payments and bonus programs. Credit cards facilitate easy electronic payments for purchases that cannot be completed with cash or checks especially online (Taherdoost, 2023). Credit cards provide flexibility through their feature which allows users to borrow funds up to a predetermined limit before repaying them later. Customers now find credit cards more enticing because of the rewards programs which include travel mileage and cashback benefits.

The operations of the retail industry together with travel and hospitality now heavily rely on credit cards as payment methods. The process of booking flights and renting cars and making hotel reservations depends entirely on the use of credit cards. The global dependence on credit cards keeps increasing because they serve as essential tools for consumer purchases and business management operations (Juusola et al., 2023).

1.1.3 The Rising of Credit Card Fraud

The increasing popularity of credit cards has triggered parallel increases in cases of illegal financial operations. The main forms of credit card fraud involve identity theft from stolen personal data for account setup and card cloning that reproduces details for fake cards and phishing scams employing deceptive fake websites or email contents to steal information (Salman, 2024).

The amount of financial loss from credit card fraud activities reaches astonishing levels. The total amount of card fraud worldwide reached $32 billion and was anticipated to rise further as in Nilson Report of 2022. Fraud results in losses exceeding financial damage because victims face emotional trauma along with feelings of insecurity (Borwell et al., 2022). Sustained occurrences of fraud negatively impact business reputations thus causing customers to lose faith in the company and abandon their loyalty.

The development of advanced fraudulent tactics requires traditional fraud detection methods to adapt because existing techniques prove inadequate against current fraud patterns.

1.1.4 Limitations of Traditional Fraud Detection Systems

Fraud detection systems of the past employed rule-based systems which established transaction flags through set rules regarding abnormal spending behaviour and specific geographical points. Simple fraud detection functions within this system are successful but the approach fails to identify sophisticated or progressing deceptive actions.

The inability to adapt constitutes one major weakness of systems that utilize rules for their functionality. The operational necessity for manual system updates emerges due to fraudster adaptations of their methods. Financial institutions experience delays in detecting fraud because of this situation which makes them prone to security threats (Shams et al., 2021). Strict transaction control rules lead to unmerited flags about legitimate monetary activities being labelled as fraudulent by the system. The system's performance becomes compromised because strict rules create misidentified transactions which both annoy customers and weaken their trust in the platform.

Traditional methods have demonstrated clear failures according to high-profile example cases. For example, systems at Target failed to recognize fraudulent activity in 2013 although warnings existed which proved the necessity of adaptable fraud prevention methods (Kolevski et al., 2022).

1.1.5 Emergence of Machine Learning

The development of machine learning (ML) represents an efficient approach for handling the weaknesses present in standard fraud detection methods. Rule-based methods differ from ML because this technology draws knowledge from past records to upgrade its ability to identify evolving fraudulent patterns (Njoku et al., 2024). ML algorithms use large transaction datasets to spot hidden patterns together with anomalous behaviour which human and traditional methods typically overlook.

ML demonstrates clear advantages in its operations. The system reviews multiple millions of transactions at fast speeds to detect complex forms of fraud with superior accuracy levels. The ability of ML models to learn independently from new information enables them to function without requiring human assistance. The fraud prevention systems decrease false alarm rates which leads to better customer satisfaction through fewer interruptions of real transactions.

Multiple financial institutions alongside payment providers currently apply ML-based systems for detecting fraud activities. PayPal implements ML capabilities to evaluate transactions right away thus enhancing its capacity to detect fraudulent activities (Aaron et al., 2024). The earliest evidence shows that artificial intelligence-based solutions identify new types of fraud before traditional rule-based systems while achieving superior accuracy which makes them a better fraud prevention tool.

The detection of emerging fraud patterns receives significant benefit from ML supervised learning methods (Random Forests and XGBoost) and unsupervised learning model applications. The continuous analysis of data through ML enables better detection rates which strengthens security for credit card transactions.

1.2 Problem Statement

1.2.1 Key Issues in Fraud Detection

The monitoring of credit card transactions for fraudulent activity becomes steadily harder because of quick changes in illegal card usage methods. The criminal models behind fraud constantly update and create new methods which defeat existing security networks. Traditional systems remain unable to detect "zero-day" frauds because these new forms of fraudulent activities escaped detection since their first

appearance (Ndungu, 2021). The use of historical data by detection models makes it impossible for them to identify fresh undetected fraud patterns. A detection system needs to evolve dynamically to identify fresh threats within real-time because fraud tactics continuously transform.

1.2.2 Dataset Challenges

The main drawback in fraud detection systems stems from their unbalanced nature. The frequency of fraudulent transactions remains extremely low relative to the total number of transactions which occur therefore they are statistically uncommon. The connection between major and minority classes results in an unbalanced model which favours predictions for legitimate transactions instead of detecting fraudulent actions (Singla et al., 2021). The incorrect identification of fraudulent transactions by fraud detection models becomes a problem since it reduces their effectiveness while also decreasing detection rates. The detection of fraud depends on proper handling of the class imbalance by using oversampling or synthetic data generation methods to provide adequate representation of fraudulent transactions (Ghaleb et al., 2023).

1.2.3 Impact of False Positives

The detection of genuine transactions as fraudulent presents a major problem that occurs through false positive identifications. False positives in the system create problems that bother customers while destroying their faith in the platform. A denied legitimate transaction disrupts customer purchasing ability which leads to frustrated customers. System performance suffers when users receive too many false alarms since it causes them to doubt the security system and ultimately choose to leave the platform. The decline in customer trust produces permanent damage to financial institution reputations which might reduce loyal customer percentages and generate revenue decreases (Van der Cruijsen et al., 2023).

1.2.4 Need for Adaptive, Intelligent Solutions

Modern fraud detection systems should be developed as adaptive intelligent systems because the methods employed in fraud schemes are continuously changing. Rule-based fraud detection systems demonstrate inadequacy because they fail to adapt to upcoming fraudulent activities. Systems need to employ real-time intelligence which can learn from data streams perpetually and change when new threats appear. Such systems need to strike the right balance between prevention of fraud acts and user-friendly experiences by preventing misleading alerts that lead to customer interruptions. Such system development would lead to enhanced fraud detection precision and higher customer trust alongside satisfaction levels.

1.3 Aim of the Study

The main focus of this research is the creation and assessment of an effective ML based system which detects credit card fraud.

- The model will benefit from training data balance through implementation of Synthetic Minority Over-sampling Technique (SMOTE) alongside under-sampling to enhance detection of fraudulent transactions.
- To executing multiple ML algorithms to determine their performance when detecting fraud. A set of four models represents the selection which includes Logistic Regression alongside Random Forest along with XGBoost and Neural Networks.
- To enhance the chosen model through hyperparameter tuning to achieve higher detection accuracy and minimize both incorrect positive and negative results.

1.4 Research Questions

The research asks three directions to determine how well ML algorithms perform while also investigating pre-processing effects together with system performance efficiency. The questions investigate the essential elements required to create a successful fraud detection system.

- how well do different algorithms detect fraud?
- what pre-processing techniques can improve the performance of fraud detection models, particularly in handling imbalanced datasets?
- which model provides the best trade-off between fraud detection accuracy and computational efficiency?

1.5 Significance of the Study

This study delivers multifaceted importance which affects both academia and practice and society and technology. The development of a machine learning-based credit card fraud detection system in this research tackles essential problems and enhances global fraud detection system effectiveness.

1.5.1 Academic Significance

The investigation adds knowledge to current studies about ML applications in fraud detection especially within credit card transaction environments. The implementation of ML algorithms for fraud detection is widespread but academic research for addressing dataset imbalance problems and changing fraudulent behaviours remains scarce (Ali et al., 2022). This study addresses data imbalance problems through

SMOTE and under-sampling applications and explores model adaptation techniques for new fraud patterns to enhance academic knowledge regarding the field. The research presents relevant discoveries about different ML methods and their implementation outcomes within authentic fraud identification systems.

1.5.2 Practical Significance

The research results provide financial organizations with practical information which helps them design more efficient fraud detection systems. Through the research, a strong ML framework would be developed which banks can use to enhance their contemporary systems for real-time fraud discovery. The implementation of this system leads to substantial reduction of financial losses suffered by institutions because of fraudulent activities that cause billions of dollars annually. Better accuracy in fraud detection models enables organizations to reduce the frequency of flagging honest customers as fraudsters and causing them unnecessary inconvenience (Sharma & Sehgal 2024).

1.5.3 Social Significance

The research protects consumers from financial fraud by taking an important social role in protecting them against increasing fraudulent practices. Research advances in fraudulent transaction identification leads to better consumer protection against financial losses as well as identity theft incidents. Such improvements increase consumer confidence because they generate greater trust in digital payment networks. The security and integrity of digital payment systems matter because the increasing number of consumers who use digital transactions needs to trust their online payments (Rasistia & Sayyidah 2021).

1.5.4 Technological Significance

Technologically, this study showcases the potential of ML in high-stakes, real-world environments, particularly in the financial sector. Digital payment expansion requires automated and intelligent systems to detect fraud because the need for such solutions has reached its peak. ML demonstrates its successful application in fraud detection by solving complex problems according to this research which illustrates how advanced technology enhances security systems. This research presentation of effective ML applications in fraud detection creates opportunities for development of future financial security technologies that use AI.

1.6 Scope of the Study

This study establishes its research parameters concerning data and methodology along with analytical procedures and necessary limitations for the investigation.

1.6.1 Data Scope

The research draws its data from the Kaggle anonymized credit card transaction dataset that provides transaction records along with labelled data about fraudulence or legitimacy. The protected personal data remains anonymous throughout the dataset while maintaining sufficient information to support ML model development and evaluation.

1.6.2 Methodological Scope

Supervised learning methods will be employed because the model needs to be trained by using data sets that distinguish fraudulent from non-fraudulent transactions. The Synthetic Minority Over-Sampling Technique (SMOTE) will be used to balance the dataset by creating artificial fraudulent transaction records because fraudulent transactions occur infrequently compared to legitimate ones. The project utilizes Logistic Regression together with Random Forest and XGBoost and Neural Networks as algorithms.

1.6.3 Analytical Scope

The analytical approach will evaluate ML model performance by using standard evaluation metrics (Rainio et al., 2024). The assessment relies on accuracy together with precision and recall and F1-score along with the area under the receiver operating characteristic curve (AUC). The selected metrics will enable a complete assessment of model success in fighting fraudulent transactions and controlling false positive identification.

1.6.4 Limitations

This research explores binary classification which means it will identify either fraud or non-fraud actions. The method operates using structured transaction information only. This research excludes analysis of unstructured data which includes text data and external fraud detection obtained from non-transactional sources.

1.7 Structure of the Study

The research presents its material in five distinct chapters which focus on different aspects of the research process.

The study begins with Chapter 1 that is Introduction to present its historical context and research objectives together with problem statement and research value. The study base establishes its essential starting point and research objectives and questions.

The Literature Review chapter of this study evaluates research about ML techniques used for fraud detection. An assessment of existing work in the literature base demonstrates both existing practices and identified research challenges and gaps that guide the evaluation of the study.

The research design with its dataset along with preprocessing procedures for missing values management and class imbalance correction stands in Chapter 3, Methodology. The development and evaluation process of the fraud detection system receives detailed explanation in this chapter.

This chapter shows the results obtained from ML models while explaining their outcome performance together with the analysis of key findings.

The study finishes with Chapter 5 which provides concluding remarks about the research results and future research directions.

Chapter 2 Literature Review

2.1 Chapter Introduction

Due to the rapid growth of digital transactions and the widespread usage of the payment systems by means of a credit card, financial fraud is becoming a frequent phenomenon in modern economical systems (Beju & Făt, 2023). Given billions of daily online and physical transactions, credit card fraud has progressed to be more sophisticated to exploit the shortcomings of conventional fraud detection tools. When it comes to finding innovative fraud techniques, these conventional weapons are a set of static and often rule based systems that are outdated. Therefore, there is a shift towards ML(ML) based approach which provides the capabilities such as adapting and being intelligent in identifying the fraudulent activities in real time (Bello et al., 2024). There is an emerging consensus in the literature that supervised and unsupervised learning models have the potential of detecting even complex fraud patterns in large and imbalanced dataset like those seen in financial transactions. Therefore, to advance fraud detection systems that are both accurate and efficient requires an understanding of the effectiveness, limitations, and comparison of strengths amongst these ML methods.

In this chapter, credit card fraud detection using ML has been explored from different research literature. First, it examines the history of fraud detection methods, a thorough description of the supervised, unsupervised, and the deep learning models. It also considers challenges in critical areas of fraud detection such as data imbalance, class mislabelling and selection of evaluation metrics. Real world applications and case studies are analysed to show how these technologies get implemented in the practice by the organisations. The chapter also examines ethical and legal issues impacting data privacy and algorithmic fairness. It finally identifies the existing gap of knowledge and introduces the conceptual framework of the present study.

2.2 Evolution of Credit Card Fraud Detection Systems

2.2.1 Rule-Based Systems and Their Limitations

In the initial attempts to catch credit card fraud, rule-based systems were heavily used in which expert defined conditions were programmed to identify suspecting transactions (Kumar & Saxena, 2022). For instance, these rules might entail unusual spending amounts, transactions outside the usual location for a cardholder, or rapid successive purchases. However, these systems worked in a rigid way and were effective only for simple fraud patterns and required continuous manual update. Fraud tactics that were not covered by rules would be allowed to go undetected. In addition, these systems created a large

number of false positives, incorrectly identifying actual transactions as potential risks, interfering with the customer experience and eroding trust between the users and financial institutions (Patel, 2023).

Another major drawback was the incapability of the rule-based systems in adapting for evolving fraud scheme. Static thresholds became insufficient as fraudsters discovered system gaps to abuse and behaviours sometimes weren't being recognised due to emergence (Ahmed et al., 2021). Additionally, the rule development was manual, causing delay in response time and thereby resulting in fraudulent transactions not being stopped in time. The use of credit cards grew especially through e-commerce and cross border payments and with that their shortcomings became obvious. However, institutions started looking out for more dynamic and automated approaches that learn and evolve without constant human involvement.

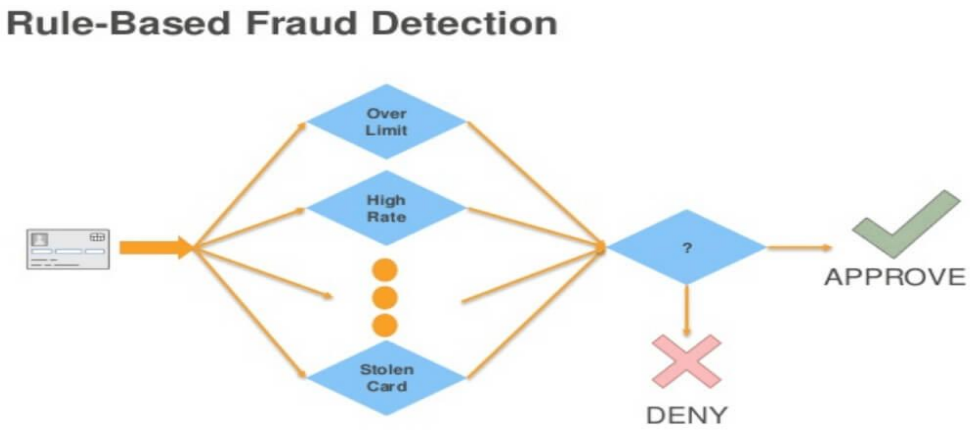

Figure 1: Rule based fraud detection approach.

2.2.2 The Shift Toward Intelligent Detection Systems

As the credit card fraud gets more complex, the industry slowly started moving towards intelligent detection systems utilising data-driven methodologies. Historical transaction data, statistical and probabilistic modeling, as well as analysis were used in these systems to discover patterns that could betray fraudulent behaviour (Mitchell, 2023). These models differed from rule-based frameworks by their ability to learn from previous data to adjust automatically for new patterns without programmer intervention. Through this capability organizations could shift from fraud detection systems built for reaction to systems that now perform predictive and proactive analysis based on learning from behaviour.

Advanced capabilities brought by ML systems intensified this development process. ML algorithms used large-scale dataset analysis for ongoing training which empowered fraud detection systems to separate valid transactions from fraudulent ones while delivering higher accuracy rates (Onwuchekwa et al., 2023).

These models solved important fraud detection problems by managing class imbalance and providing real-time response capabilities. Financial institutions shifted toward intelligent scalable systems which maintained fraud detection capability when new emerging fraud methods appeared.

2.3 Traditional Fraud Detection Techniques

2.3.1 Manual Review

Early digital banking periods together with new credit card development witnessed complete manual transaction analysis by skilled staff who worked to detect fraud (Hilal et al., 2023). Teams of fraud analysts in financial institutions were employed and teams examined the transaction data using domain expertise and customer profiles to determine which transactions seemed fraudulent. The analysts would have flagged purchases that were unusual in terms of size, in an abnormal geographic location, or markedly out of line with that customer's purchase patterns. It gave a personalised and context aware detection, worked best for the low volume of transaction environment. The manual method had become inefficient as the transaction volumes rose, in parallel with the rise in the use of credit cards and e-commerce (Porwal & Mukund, 2019). Although it was designed to handle real time financial activity, it did not have the speed and scalability necessary to support it. Moreover, the manual reviews were reactive, using to identify fraud only after the fraud has taken place, reducing their preventive capabilities. Frequent and more complex fraud resulted in a wide margin of human error and delay which revealed the shortcoming of analyst intuition.

2.3.2 Rule-Based Systems

The implementation of rule-based systems created a fundamental basis for automation through their ability to enhance both the consistency and speed of fraud detection. The systems processed data through sets of predetermined if-then rules. Rules within the system monitor high-value transactions and purchases made beyond typical purchasing areas. User behaviour analysis through binary logic systems did not detect the subtle differences in activity contexts. Rule-based systems provided benefits yet faced several severe limitations in their operation (Khurana, 2020). The systems demonstrated inflexibility by being unable to adapt to the changing fraudster methods. When new fraud patterns entered, the analysts needed to develop manual detection methods followed by new rules which led to system lag and outdated processes. Their rigid binary storage systems did not accurately capture the multi-faceted user situation contexts. Customers end up feeling frustrated because their payments get rejected and accounts get blocked from acceptable spending activity detected as fraudulent (Mir, 2024).

Technical standards in fraud detection experienced problems due to their tendency to produce numerous incorrect positive results. The screening process for plaintiffs operated when legitimate transactions incorrectly obtained fraud labelling. High threshold limits used to stop fraudulent activities create unnecessary blocks for valid users. Flexible rules compromised the system effectiveness by allowing actual fraudulent transactions to go undetected. The system suffered from deficiencies due to its inability to evaluate various fraud patterns simultaneously. The system failed to detect advanced forms of fraud because it did not have the capability to analyse multiple data points simultaneously. These rigid transaction systems failed to adapt to the faster pace of development in payments and fraud schemes that transformed transactions (Olushola & Mart, 2024). Basic fraud detection solutions showed their limitations for high-frequency rapid systems because of their solid structure combined with delayed detection capabilities.

2.4 Rise of ML in Fraud Detection

2.4.1 ***ML vs Traditional Methods***

Fraud detection has experienced fundamental changes through the implementation of ML which solves existing system weaknesses (Madhurya et al., 2022). The manual threshold limit management of Rule-based systems also differs from ML models that directly learn from data to evolve with experience. Traditional systems maintain rigid programs that struggle to uncover new fraudulent approaches which do not meet their defined rules. The ability of ML algorithms to identify slight changes in transaction patterns allows them to detect more accurate cases of suspicious activity. The system can process large transaction volumes at real-time speeds which lets them adapt to growing financial environments.

The primary advantage results from reduced instances of false alarm detection (Trivedi et al., 2020). Inflexible transaction rules of traditional methods trigger many valid transactions which leads customers to become dissatisfied. The combination of time, location, frequency data and spending patterns helps ML models create more intelligent decisions. The data-driven adaptive approach of ML systems provides both better precision and recall in their decisions. Financial institutions utilise ML-based technology to protect assets and build trust as they work to deliver efficient operations while combatting fraud. Intelligent systems using adaptive logic have replaced traditional fixed-rule systems to provide the most significant improvement in credit card fraud detection security.

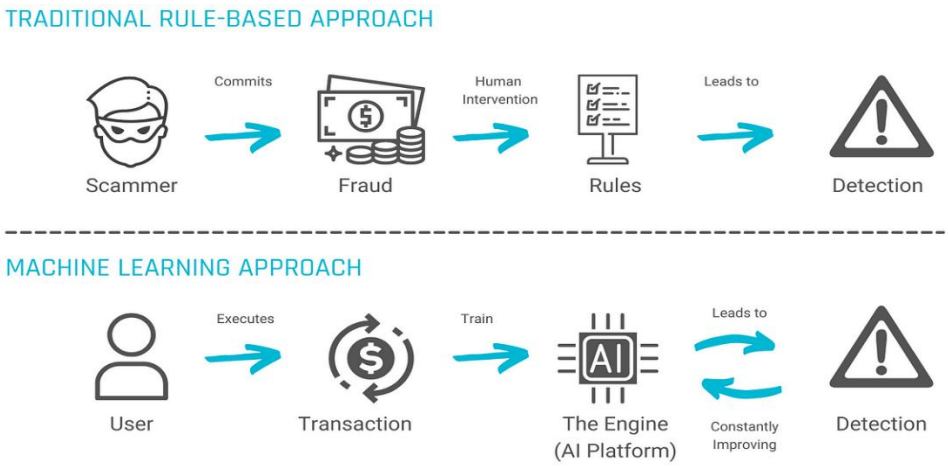

Figure 2: ML vs Rule based approach. (Matsuk, 2022)

*2.4.2 **Common ML Algorithms***

Multiple ML algorithms bring specific advantages to credit card fraud detection while working with different datasets and detection purposes. Logistic Regression functions as a common baseline model because it efficiently performs binary classification while enabling linear interpretability of the model output (Geng et al., 2024). Linear relationships between features and outcomes make the model perform effectively. Random Forest and XGBoost under the Decision Tree-based category have attained universal recognition due to their strong accuracy and their resistance to outliers as well as their effectiveness in identifying nonlinear patterns.

Random Forest uses multiple decision tree models to combine their results into generalised predictions that overcome fitting issues. XGBoost implements sequential error-correction modeling which boosts its predictive strength when detecting unbalanced data fraud patterns (Cruz et al., 2024). Deep Learning models and Neural Networks gain momentum in the market due to their power to recognise hidden patterns as well as model complex variable interactions. ML models perform well on big data but their processing requirements force users to fine-tune their operation parameters. ML in fraud detection excels because its technology learns and adjusts its procedures automatically as events unfold in real time.

2.4.3 Real-Time Adaptation

Real-time adaptation stands as the greatest advantage when using ML to detect fraud. Different from static platforms, ML models process new information to learn underlying trends which result in adjustments of operational parameters for detecting new fraud patterns (Bello et al., 2024). Real-time

learning plays an important role in domains with rapidly changing fraudulent methods that cannot be predicted. Real-time system performance increases notably when algorithms merge with reinforcement learning techniques.

The models create better forecasts when provided with recent decision feedback data. The predictive models receive recent decisive outcomes to improve their accuracy in forecasting results. The system learns from false flags on legitimate transactions to prevent repeating such errors in future cases (Afriyie et al., 2023). Improved adaptive capabilities enable the system to recognise suspicious activity at unknown fraud levels including zero-day attacks effectively. Organisations face multiple difficulties during the implementation of ML-based fraud detection systems which need proper handling.

2.4.4 Challenges in ML Detection

Multiple essential challenges stand in the way of successful ML-based fraud detection system deployment. The proportion of fraudulent transactions remains relatively low compared to total transactions which creates a training challenge for many traditional algorithms. The fragmentation of data harms model functionality while consistency remains a fundamental requirement. A model becomes overfitted after it specializes too much to training data triggering its inability to work properly (Nuthalapati, 2023).

Another concern is model interpretability. Black box models including deep neural networks present challenges for analysts who aim to understand their decision-making processes (Mienye, & Jere, 2024). The inability to track prediction processes creates concerns in regulatory systems because they need detailed explanations while complying with requirements. The expense of computation stands as a significant issue mostly affecting organisations that possess minimal computing resources. The requirements for real-time fraud detection which include low latency and high processing power limit the compatibility with some ML models. When models find detailed patterns in their training sets this produces overfitting which makes them unable to process new inputs correctly. Assessing fraud requires pre-trained algorithms that work with datasets which include explicit labels differentiating between fraudulent and non-fraudulent cases.

2.5 Supervised Learning Models

The model utilises known data relationships in past scenarios to establish prediction capabilities for future unknown scenarios. The algorithms learn from training datasets which contain labelling between fraudulent and non-fraudulent target information (Khatri et al., 2020). The model extracts input patterns together with their corresponding output labels to make predictions about new observations. The

widespread binary classification model called logistic regression helps detect fraudulent activities. Prospectively supervised systems prevail in broad applications because of their capability to match data profiles together with set system requirements.

2.5.1 Logistic Regression

The identification of binary classes for fraud detection through analysis uses logistic regression as a fundamental and widely used method among operators. Through logistic functions it calculates predictive probabilities to determine which category, fraudulent or legitimate a transaction belongs to (Wang & Zhao, 2022). The simple nature of logistic regression detects linear patterns effectively in the data which makes it popular among fraud detection systems. Decision trees succeed in fraud detection by providing instant results while retaining a straightforward output visualisation. Random Forest utilises several decision trees which use different features from random subsets of data to generate predictions by aggregating responses through voting mechanisms.

2.5.2 Decision Trees and Random Forest

The valuable characteristic of decision trees lies in their ability to clarify predictions by defining branches according to feature criteria which lead to classifications. Decision trees excel at fraud detection because they offer quick calculation times together with simple understandable results. Single decision trees experience poor generalisability because they tend to overfit training data therefore they perform poorly when identifying new unseen data. Random Forest ensemble techniques help address this limitation when used in applications. Random Forest constructs many decision trees from divided data subsets through which it combines structural predictions using majority voting (Dileep et al., 2021). This system uses distributed processing while implementing regularised methods to prevent overfitting for extensive applications. Random Forests become ideal for identifying fraud within high-dimensional datasets because they recognise complex patterns using nonlinear interaction methods without requiring extensive data manipulation.

2.5.3 XGBoost

XGBoost delivers an effective and optimised version of gradient boosting algorithms. Each successive model built through this system strives to fix mistakes from the previous stage. Systems integrating XGBoost now serve as standard fraud detection tools for processing datasets that are imbalanced alongside handling both missing values and outliers (Kabane, 2024). Application of SVMs requires users

to optimise hyperparameters and select kernels which leads to increased computational demands when working with large datasets. The built-in analysis tools in the model help analysts determine which variables drive fraud classification most strongly. The skills of SVMs include detecting outliers while maintaining scalability yet their dimensional constraints make them suitable for modest fraud datasets which require limited outlier sensitivity and scalability.

2.5.4 Support Vector Machines (SVM)

SVM represent a robust classification method which searches for an optimal hyperplane between different classes of data points during execution (Kumar et al., 2022). The detection of fraud becomes more effective with SVMs due to their ability to operate in high-dimensional spaces alongside their proficiency in detecting fraud when crimes occur rarely. SVMs gain the ability to detect complex fraud patterns through non-linear decision boundaries created by kernel functions. Neural networks need powerful computing machines for training purposes combined with long processing periods and deliver outputs based on the quality of their training datasets. SVMs show limitations in handling large datasets in real-time and are sensitive to outliers and require smaller datasets to reach precise results. They are particularly effective in smaller curated datasets and fraud applications.

2.5.5 Neural Networks

The supervised learning modeling framework delivers multiple detection tools for credit card fraud but demands consideration of specific advantages and considerations between models. The structure of these artificial intelligence models contains several layers made of interconnected processing nodes known as neurons which are connected through weighted relationships. Network models analyse voluminous datasets to discover hidden patterns that basic systems cannot detect yet their limited transparency makes them unsuitable for financial prediction work because financial models always need clear explanations. Neural network training duration and needed computational power directly correlate with the variables used for training data quality yet trainability (Georgieva et al., 2019). By using experts in pre-processing and feature scaling together with regularisation techniques neural networks produce promising outcomes in detecting fraud.

The multiple analytics tools from supervised learning models have different implementation requirements to detect credit card fraud. The selection of model depends on how the system needs to function because it must either achieve maximum accuracy or understanding or perform quickly or supply flexible

resources. The combination of multiple practical models occurs in ensemble methods or layered model architectures to achieve best performance results.

2.6 Unsupervised and Semi-Supervised Techniques

2.6.1 Clustering (K-Means)

Using K-Means clustering techniques allows businesses to group transactions based on similarities because it does not require labelled transaction data. Clusters help fraud detection teams identify patterns which differ from typical consumer conduct. The distinct behavioural patterns between fraudulent and ordinary transactions allow both clusters to form separate entities or remain separate from one another. The performance of clustering algorithms decreases when the data contains uneven groups since they only work well with homogenous clusters. K-Means operates at a fast execution speed yet needs cluster count input from the user although real-world fraud patterns might not always match this pre-set value (Jain et al., 2021).

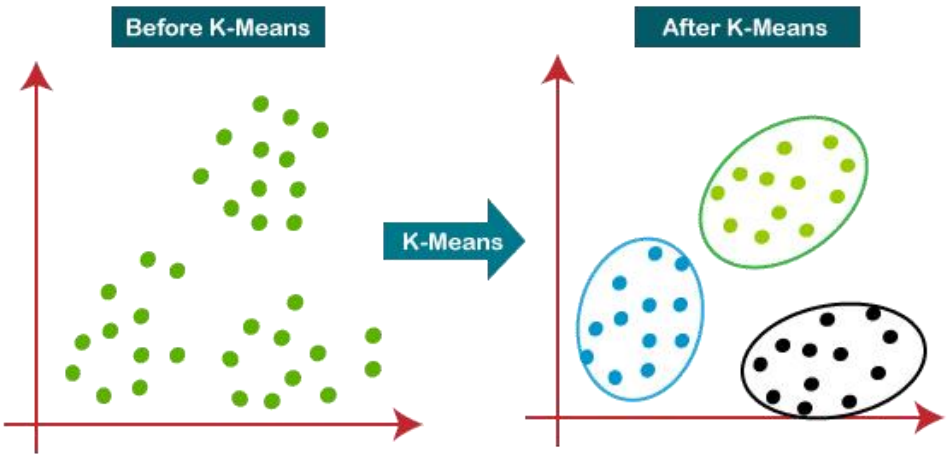

Figure 3: K-Means clustering applied to an unlabeled dataset (Jain, 2023)

2.6.2 One-Class SVM

One-Class Support Vector Machines excel at fraud detection tasks because they operate effectively even when fraudulent data examples are rare (Leevy et al., 2023). The models receive training solely from genuine transactions to learn standard user behaviour. A major difference identified by the model from this established pattern will automatically trigger the anomaly detection system. One-Class SVM shows high effectiveness in high-dimensional applications and does not need labelled fraudulent examples for training. The model produces false detections when normal data contains excessive noise or outliers or

legitimate activity shows significant variability. This method performs reasonably well as an initial approach to fraud detection when fraud labelling occurs infrequently.

2.6.3 Isolation Forest

The anomaly detection technique Isolation Forest separates observations by performing random feature selection and value splitting operations (Rajeev & Devi, 2022). Anomalies become detectable easier because they stand apart from general patterns through their scarce occurrence. The anomaly detection method proves useful for identifying unusual fraudulent activities in fraud detection systems that do not need labelled transaction data. Organisations use these techniques for fraud detection by learning patterns from current dataset activity instead of traditional historical label systems. This model calculates anomaly scores for each transaction to help investigators determine their investigation sequence. Performance of the model decreased when deceptive input data followed regular behavioural patterns.

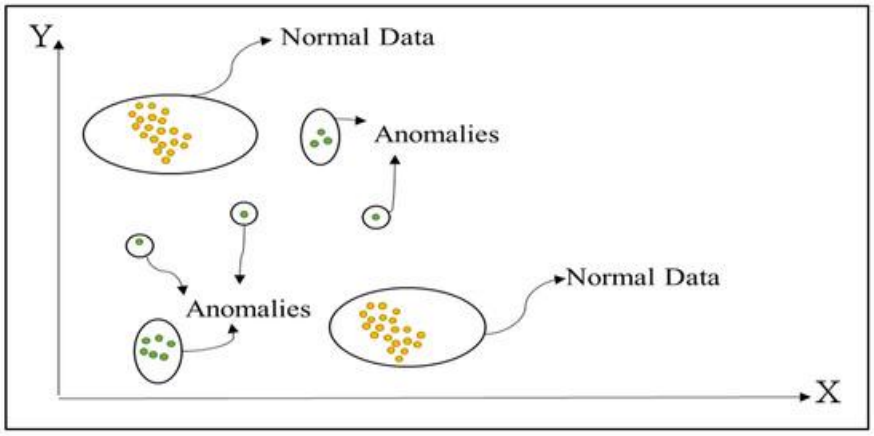

Figure 4: Anomaly detection, separating normal data clusters from isolated anomalies. (Maram Alamri & Mourad Ykhlef, 2022)

2.6.4 Use Cases in Limited-Label Scenarios

Real-life fraud detection frameworks need successful implementations of self-directed and partly supervised techniques because they demonstrate effective results when working with data having limited labelling accuracy or currency challenges. The process of pattern extraction from current data enables organisations to detect new fraud patterns regardless of limited historical labelling data. Supervised learning systems work together with financial institutions to deploy these models as part of their anomaly detection process. The flexibility of ANNs allows users to adapt their models between binary classification and anomaly detection needs within various detection systems (Plakandaras et al., 2022).

These flawed methods provide essential functionality to organisations that lack sufficient labelled data and face rapid changes in their fraudulent patterns.

2.7 Deep Learning Approaches

2.7.1 Artificial Neural Networks (ANN)

ANNs represent computational structures which derive their design inspiration from a human brain pattern called the neural structure. The system builds layered structures from nodes to process input characteristics by applying cumulative addition rules followed by activation rules. ANN operate within credit card fraud systems because they recognise non-linear data connections in transaction data relationships. Such models prove most effective for large data sets because they discover complex patterns which standard models cannot identify. General ML models known as ANNs enable anomaly detection and binary classification applications along with multiple security detection methods (Kilickaya, 2024). The application of ANNs depends on powerful computing equipment and precise parameter handling to prevent overfitting results. Such models operate in self-contained environments making it hard for people to understand them yet regulated financial institutions need clear explanations for their decisions. The main reason why ANNs find application in production systems is their capability to achieve high prediction accuracy despite lacking transparency in interpretations. These networks enhance fraud detection value by analysing schemes which advance through multiple steps instead of single incidents.

2.7.2 Recurrent Neural Networks (RNN), LSTM

The processing of sequential time-dependent patterns is best achieved through deep learning systems including RNNs and LSTMs. Similarly, to other deep learning models RNNs along with LSTMs face obstacles regarding interpretability which institutions with transparent practices may find challenging to adopt. As an unsupervised neural network model autoencoders perform anomalous pattern detection while reducing data dimensions (Rezapour, 2019). Through their solution to standard RNN vanishing gradient problems LSTMs enhance their ability to detect long-term dependencies in data. Through training an LSTM model, it can determine when a recent transaction exceeds established patterns of user behaviour spanning the previous week or month. These networks excel at finding criminal activities that build over multiple transactions rather than appearing suddenly. Minimal weight along with built-in flexibility in the network design makes it efficient in detecting core behavioural patterns from normal data. Autoencoders have established themselves as an effective tool for detecting fraud by utilising their ability to improve

existing security frameworks while recognising intricate patterns of attacks (Fanai & Abbasimehr, 2023). The sequential pattern modeling features of these tools enables their use as strong tools for analysing fraud through behavioural patterns.

2.7.3 Autoencoders for Anomaly Detection

The unsupervised artificial neural network system known as autoencoders performs dual functions by reducing data dimensions and detecting anomalies. Such models will first convert the input data into a reduced dimension space before it can project the original input from this low dimension space. Autoencoders training so as to reproduce normal transactions successfully creates minimal reconstruction error. When it has high reconstruction errors, a model detects fraudulent transactions because it lacks reconstructing data points far from its learnt data distribution. Autoencoders provide outstanding benefits to the fraud detection systems since it uses an unsupervised learning approach of working with unlabelled data to detect new fraudulent activities at full independence. The benefits of the neural networks include their lightweight design that provides both operational flexibilities and standard detection capability. Dependence on this model is entirely subjected to training data quality since as such conditions may result in misidentifications of intricate transactions that resemble normality. Subjective thresholds which are adaptable form the basis of anomalous event detection and display various values in separate datasets. Autoencoders remain popular because they allow the more traditional fraud detection algorithms to reveal hidden attack patterns without breaking the system constraints (Zou et al., 2019).

2.8 Hybrid and Ensemble Techniques

2.8.1 Bagging Techniques

Bootstrap Aggregating functions under the ensemble learning frame reduces model variance and reduces the overfitting problem (Ngo et al., 2023). Decision trees are the primary approach for bagging within fraud detection systems, training different models against random portions of training data. The model obtains both accuracy enhancement and stability increase through a majority vote process applied to its outputs. Random Forest serves as a unique illustration through its ability to create multiple diverse decision trees that perform effectively in noisy and imbalanced credit card fraud scenarios. Bagging proves useful in situations requiring robustness and data distribution generalisation capabilities. The parallel design of this learning algorithm enables efficient real-time processing of high-volume data making it appropriate for modern fraud detection systems (Khalid et al., 2024).

2.8.2 Boosting Techniques

The predictive model accuracy from boosting algorithms develops through continuous training of weak learners targeting the errors produced by earlier models (Mienye & Sun, 2022). The ability of boosting algorithms to improve performance in rare fraud detection situations makes them valuable for fraudulent transaction detection. Boosting methods XGBoost and AdaBoost remain widespread across many applications. XGBoost stands out as the preferred boosting method because it features regularisation capabilities alongside exceptional data processing speed when processing structured data (Mohbey et al., 2022). Each boosting iteration gives increased importance to misclassified examples so the model optimises its capacity to handle challenging transactions. Their ability to detect subtle fraud patterns surpasses traditional models because of their effectiveness in detecting such patterns. The strength of boosting algorithms requires professional parameter tuning to maintain appropriate data fit during operations in noisy environments.

2.8.3 Stacking Models

Stacked generalisation which operates as stacking functions as an advanced ensemble method for combining base model predictions through a meta-model approach. Stacking differs from bagging or boosting methods as it merges numerous heterogeneous algorithms such as logistic regression with decision trees and support vector machines to optimise predictive success. A meta-model learns from base models' outputs to detect which algorithms produce the most successful results within certain conditions. Fraud detection performance increases through stacking because it allows integration of diverse algorithmic strengths (Abbasi & Shah, 2022). A single model highlights high-value fraud cases whereas a different model focuses on transaction timing anomalies. The layered design approach enhances threat assessment by providing extensive insight into potential risks which makes stacking prove beneficial for advanced fraud detection systems.

2.8.4 Hybrid Approaches

Enhanced fraud detection occurs through hybrid models which unite supervised learning alongside unsupervised learning to extract maximum detection capability (Carcillo et al., 2021). The models work best for scenarios which face restricted labelled data resources or fast-changing fraud patterns. A hybrid approach combines XGBoost supervised algorithms with Isolation Forest unsupervised anomaly detection algorithms so systems can discover familiar fraud patterns alongside new unknown patterns. The supervised model teaches itself using previous historical fraud data and simultaneously the unsupervised element finds suspicious patterns outside established standards. The implementation of two parallel models creates a framework that demonstrates enhanced defence mechanisms against emerging

fraudulent practices. Financial institutions get their best results by implementing hybrid systems as part of multi-layered detection frameworks.

2.9 Data Challenges in Credit Card Fraud Detection

2.9.1 Class Imbalance

Detecting fraud presents a major challenge because fraudulent transactions occur rarely within extremely large datasets. Real-world data often presents less fraudulent activities making it difficult for ML models to discover discriminatory patterns (Makki et al., 2019). Standard classifiers develop an unfair bias toward non-fraud transactions which produces high accuracy results but fails to discover rare fraudulent instances. The skewed distribution damages evaluation metrics and compromises the model's practical application value. The models struggle to detect actual fraud when their continual prediction of transaction legitimacy becomes a persistent issue. The development of minority class representation methods remains essential both to improve fraud detection accuracy and decrease false negative rates in operational settings.

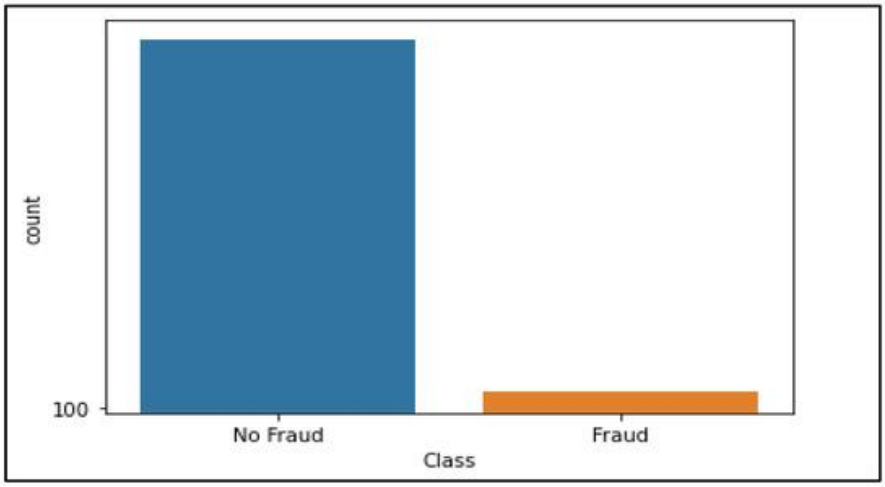

Figure 5: Class Imbalance between non-fraudulent and fraudulent transactions (Alamri & Ykhlef, 2022)

2.9.2 Resampling Techniques

The application of resampling techniques serves as a standard method to balance datasets before classifier training runs. Two oversampling methods called Synthetic Minority Over-sampling Technique (SMOTE) and Adaptive Synthetic Sampling (ADASYN) create artificial examples of minority classes to boost their presence in datasets. The minority class instances number can be matched by under-sampling which reduces the majority class instances. The two methods strive to construct balanced training data that

enhances a classifier's ability to identify fraudulent activities. The combination of oversampling techniques with SMOTE and ADASYN results in overfitting when synthetic data looks too similar to genuine cases whereas under-sampling eliminates beneficial data points (Brandt & Lanzén, 2021). Recently developed hybrid approaches leverage both strategies but need proper implementation to prevent the damage of model generalisation abilities and reduce unwanted training noise.

2.9.3 Label Noise and Misclassification

Fraud datasets often struggle with label noise because transaction classifying errors produce mislabelling between fraudulent and legitimate deals (Zhang et al., 2021). The practical implementation shows some fraudulent transactions rise through the system reporting as legitimate while some authentic transactions trigger false fraud alarms. Model training suffers due to misclassification in datasets because supervised learning algorithms require precise ground truth information. Manual review errors combined with reporting delays and unclear determination of fraudulent patterns lead to label noise in datasets. The integration of imperfect labels into training data leads predictive models to detect false patterns which increase rates of incorrectly flagged transactions. Data verification along with external feedback loops and semi-supervised approaches enable the mitigation of label noise in data sets.

2.9.4 Feature Limitations

The effectiveness of fraud detection models depends strongly on the quality range of provided input features (Chaquet-Ulldemolins et al., 2022). A model's ability to differentiate between legitimate behaviour and fraud becomes restricted when features are incomplete or sparse or lack information. The analysis of fraud detection data faces various obstacles including absent transaction details alongside inconsistent merchant labelling and data privacy protocols that blur information. The detection of evolving fraud patterns requires features that can capture behavioural and temporal dynamics which static features may not achieve. Low quality features in a model produce decline in predictive power and result in the system adopting shallow patterns which perform poorly for unknown scenarios. Domain-specific expertise together with feature engineering enables the creation of specialized features that improve model performance through methods such as average transaction interval calculation and user behaviour deviation determination.

2.9.5 Temporal and Real-Time Constraints

The prevention of unauthorised transactions depends on real-time operation from modern fraud detection systems. Migration of fraud detection systems to real-time operation introduces multiple performance challenges alongside time drift problems and challenges when processing streaming data (Bello et al., 2023). When fraud patterns transform over time the predictive models trained from previous data can become ineffective unless model updates are performed regularly. The retraining of fraud detection models with new cases becomes impeded because of slow data labelling procedures. Real-time systems demand optimised infrastructure alongside precise algorithms that execute quick decisions with equal precision to speed. The ongoing process of model deployment combined with lifecycle management remains difficult for high-frequency financial environments because of these constraints.

2.10 Evaluation Metrics and Model Effectiveness

2.10.1 Accuracy and Its Limitations

Classification tasks normally use Accuracy for evaluation which calculates the correctly predicted instances against the total number of predictions. In credit card fraud detection operations which handle highly imbalanced datasets accuracy metrics can provide deceptive results (Mittal, S., & Tyagi, 2019). A model that indicates all transactions are legitimate could reach over 99% accuracy because fraud cases rarely occur but it would fail miserably at its essential function. An inaccurate but misleading performance metric hides the detection deficit of the model so it appears effective despite missing real fraudulent cases. Accuracy alone proves inadequate as a standalone performance metric as it fails to capture the importance of correctly detecting minority cases like fraudulent transactions. Therefore, it needs additional metrics which better measure the model's abilities within the high-risk fraudulent class.

2.10.2 Precision, Recall, and F1-Score

The quality assessment of fraud detection systems depends heavily upon two important metrics: precision and recall. Precision shows the ratio of correctly identified instances between fraudulent cases while highlighting the model's ability to prevent misleading reports. The recall measure shows the model's ability to detect fraud by counting how many actual instances it successfully identifies. The F1-score combines precision and recall through their harmonic mean because it provides results when enhancing one metric degrades performance in the other. Detection systems in fraud mitigation need to balance recall for complete fraudulent case identification with precision to reduce unnecessary false alarms. These combined metrics help develop an advanced performance evaluation system for models beyond basic accuracy metrics (Hashemi et al., 2022).

2.10.3 ROC-AUC vs. PR-AUC

ROC-AUC serves as the widespread metric for evaluating model categorization abilities. A ROC-AUC calculation performed on data with significant imbalance leads to unrealistic favorable outcomes because it evaluates true positives alongside inflated true negatives affected by data imbalance (Alkattab & Edén Wallberg, 2024). PR-AUC measures fraud detection with greater effectiveness than other methods because it uses precision and recall statistics to evaluate minority class outcomes without considering dominant class predictions. The PR-AUC metric reveals a model's ability to find rare events better than other metrics and adapts more effectively to fraud detection requirements because it values accurate fraud recognition above standard classification performance.

2.10.4 Cost of Misclassification

The incorrect classification of fraudulent transactions results in substantial financial expense alongside damage to organizational reputation. When false negatives occur by missing fraud the results lead to monetary losses and damage both customer loyalty and regulatory compliance (Mwangi, 2024). Incorrect identification of legitimate transactions as fraudulent occurrences produces customer experience difficulties and harms brand reputation while raising operational spending through manual inspection requirements. Model evaluation needs to take account of these unequal costs instead of using statistical metrics alone. Organisations that practice cost-sensitive learning assign customised error weightings to match their business expense priorities. Success in fraud detection depends on finding an appropriate equilibrium between detection effectiveness and operations efficiency. The goal should be finding a way to reduce false negative results while staying clear of excessive false positive outcomes which strain operational processes and diminish customer satisfaction.

2.10.5 Metric Selection in Practice

The financial institution's operational goals together with its risk tolerance and resource capacity determine the selection of appropriate evaluation metrics in real-world applications. Organisations which handle high-value transactions focus on recall because they need to confirm detection of all fraudulent activities despite sacrificing precision accuracy (Vaquero, 2023). The focus on precision becomes critical for institutions trying to decrease operational strain and false alert notifications. Financial institutions track their performance by combining F1-score and PR-AUC metrics to achieve balance between different operational factors. Model validation and adjustment of thresholds should be performed

regularly to maintain the relevance of selected metrics during which fraud patterns alongside business environments transform over time.

2.11 Real-World Applications and Industry Practices

2.11.1 Fraud Detection in Financial Institutions

Major financial institutions now use ML technology to improve the accuracy rate and speed and efficiency of their fraud detection operations. The traditional rules-based pattern detection methods became incapable of keeping up with complex fraud activities alongside changing customer transaction behaviours. JPMorgan Chase and HSBC use supervised and unsupervised ML models to detect anomalies in the billions of daily transactions they process every day (Nweze et al., 2024). The analysis system examines multiple variables including transaction timing and location and spending patterns and device metadata to identify suspicious activities in real time. The systems use adaptive learning that allows failure data provided by human analysts and direct customer feedback to improve model performance continuously. The adoption of predictive analytics by banks has resulted in fewer financial losses combined with reduced false alarms alongside elevated customer trust levels. The financial industry faces strict regulatory demands while needing to maintain model transparency as well as auditability and compliance standards.

2.11.2 Case Example: PayPal's Real-Time System

PayPal uses its famous AI-driven real-time fraud detection infrastructure to build its reputation for effective security capabilities (Khurana, 2020). The company processes millions of transactions each day through deep learning models which evaluate each transaction in milliseconds. Advanced models from PayPal combine historical user behaviour with device intelligence and geolocation information and velocity detection to detect fraudulent transactions. PayPal stands out through its integration of supervised learning with anomaly detection and rule-based overlays which allows the system to detect existing fraud patterns alongside new emerging threats (Aaron et al., 2024). The system develops its capabilities automatically through continuous model updates with newly labelled transaction data which enables it to react to emerging fraud patterns. PayPal uses its dual role as payment processor and digital wallet provider to track user activities across all connected platforms. The comprehensive visibility of the system enables both high detection rates and low customer hassle. PayPal's successful fraud prevention model demonstrates that real-world systems succeed by merging extensive data sources with quick adjusting algorithms for operational security.

2.11.3 Mastercard's AI-Driven Strategy

Mastercard created its enterprise-scale AI platform decision intelligence to stop fraud through advanced analytics and real-time ML with transaction detection capabilities (Ahmadi, 2023). Through analysis of historical customer data and merchant profiles and contextual information the system produces risk scores for each transaction. Instant transaction decisions are based on scores that help banks determine approval or denial. Mastercard developed a system which tackles the frequently occurring false decline problem to improve both customer satisfaction and merchant trust. The company's artificial intelligence methodology encompasses multiple stages of the customer journey from onboarding through account protection and post-transaction chargeback management. The investment into explainable AI by Mastercard allows for transparent and auditable decision-making processes. Financial partnerships and regulatory adherence depend heavily on maintaining dual emphasis between performance and compliance requirements. Mastercard demonstrates that AI implementation within real-time transaction pipelines enables fraud prevention to be highly precise without impacting operational efficiencies.

Figure 6: Companies Using ML for Fraud Detection. (Bharadwaj, 2021)

2.11.4 Regulatory and Operational Integration

Fraud detection systems in industry achieve their maximum effectiveness through precise alignment with specific regulatory frameworks and operational business foundations. At a global level financial institution alongside payment platforms must fulfil the requirements of the Payment Card Industry Data Security Standard (PCI DSS) and General Data Protection Regulation (GDPR) and their corresponding national anti-fraud directives (Razikin & Widodo, 2021). The data collection process together with storage practices and analysis functions come under specific regulatory control especially for customer-based behavioural data. Fraud system developers build their solutions to deliver high detection accuracy combined with guarantees of data privacy alongside audit capabilities and readable explanations. Real-

time transaction engines contain fraud systems that need to respond with decisions within milliseconds of operation. The implementation of these systems demands scalable architectures together with efficient model deployment frameworks and minimal-latency pipelines. Organisations achieve performance improvement through feedback loop connections between automated fraud detection systems and human review teams. The combination of regulatory alignment and smooth integration enables sustainable scale-based fraud prevention practices.

2.12 Ethical and Legal Considerations

2.12.1 Data Privacy and Protection

The analysis by fraud detection systems depends significantly on highly confidential transactional data and behavioural information while creating substantial privacy issues. These systems handle customer information and device metadata together with location data and spending histories which makes them vulnerable to data privacy protections like the GDPR and CCPA and PCI DSS. Organisations that adopt these frameworks should be data transparent by limiting their data acquisition processes and operating under rigid legal data processing rules. The use of ML algorithms adds more barriers due to the reason that trained models may contain personal identification details (PII) during processing sensitive information (Arambawela & Aponso, 2024). Organisations combine a number of anonymisation tools as well as tokenisation and encryption in order to manage PII protection processes. For organisations, reaching full compliance while achieving model accuracy is still tricky. Data privacy and utility need equilibrium management from institutions while facing dual requirements from legal compliance frameworks and public trust evaluations.

2.12.2 Fairness, Bias, and Explainability

ML systems show increased influence in financial decisions yet researchers focus more intensely on algorithmic fairness and bias concerns. Training fraud detection models with historical data can preserve systematic biases that lead to consistent misidentification of transactions based on either geographic location or user identity characteristics (Kamalaruban et al., 2024). Such bias creates inaccurate fraud alerts that discriminate groups of people marginalizing them and potentially limiting their access to financial services or damaging their reputation. Financial institutions must perform bias tests to protect their operations and must achieve balanced distributions of training data across their systems. Transparency and explainability have become vital requirements particularly in environments that need regulatory compliance. Fraud models which utilise neural networks and ensemble methods remain

complex and fundamentally non-interpretable systems. Internal validation and regulatory compliance gain assistance through SHAP and LIME tools that help reveal model prediction-influencing features (Okenwa et al., 2024).

2.13 Research Gaps in the Literature

2.13.1 Limited Use of Unstructured and Behavioral Data

Most credit card fraud detection systems base their decision-making on structured transactional data but research has mainly neglected the potential of unstructured and behavioural data. Unstructured data points including customer service logs together with browser behaviour and clickstream patterns and device metadata provide essential contextual clues for fraud detection. These data points help detect intent as well as separate random mistakes from planned malicious activity. The current fraud detection methods and datasets primarily analyse fixed attributes including transaction amount, location and merchant category. The limited range of this focus limits model efficacy particularly during complex fraud situations that include synthetic account creation and identity theft and social engineering. The academic field needs more research on Natural Language Processing (NLP) and behavioural analytics and multimodal learning topics. The lack of publicly accessible databases which connect structured and unstructured data elements restricts the advancement in this field. The research gap between human-like fraud detection models and contextual decision-making capabilities points to a significant opportunity for improved detection systems.

Figure 7: Sources of Unstructured Data for Fraud Detection ("PayPal's Use of Machine Learning to Enhance Fraud Detection (and More) - Technology and Operations Management," 2018)

2.13.2 Gaps in Real-Time and Adaptive Modeling

The critical need for real-time fraud detection in modern payment systems demands more comprehensive research into adaptive online learning models that adapt their fraud pattern recognition to emerging fraud patterns in real time. Most studies in academic institutions depend on static datasets to evaluate models offline instead of reproducing operational challenges which occur during real-time deployment. Models trained with historical data risk becoming invalid if fraud tactics shift quickly throughout a period when proper updates to the models are absent. Research demands more investigation into streaming data frameworks and low-latency feature processing together with real-time system feedback mechanisms. Emerging online ensemble learning and reinforcement learning methods hold potential for continuous learning but show minimal practical application when detecting fraud. The research field lacks extensive investigation into dynamic threshold adjustment strategies which modify decision logs based on transaction contexts and fraud occurrence frequencies. Resilient fraud detection systems need a solution to connect static evaluation approaches with dynamic deployment models to achieve effectiveness in high-frequency real-world scenarios.

2.13.3 Comparative Evaluation of Model Architectures

Multiple ML approaches including logistic regression and neural networks along with hybrid systems exist for fraud detection yet systematic comparisons of these models remain scarce. Most research produces evidence supporting particular models without conducting experiments that compare different models on identical experimental parameters. The lack of systematic comparisons means researchers cannot reliably develop findings across different models or find appropriate solutions for distinct fraud situations. XGBoost and Random Forest ensemble methods demonstrate strong performance on structured data but deep learning models show better results when applied to behavioural and sequential pattern analysis. A lack of research exists which directly evaluates these distinct approaches through standardised measures including PR-AUC alongside cost-sensitive loss functions. The trade-off between model performance and interpretability is frequently analysed via anecdotal methods instead of empirical testing protocols. Research comparing these models through accuracy and computational cost measures in addition to fairness and explainability assessments will generate practical insights that benefit academic researchers and industry practitioners. The adoption of standardised metrics would improve strategic decision-making related to fraud detection technology deployment and selection.

2.14 Chapter Summary

Through an analysis of existing scholarly work this chapter examined the development and present state of credit card fraud detection. A review showed that historical system development started with manual

and rule-based solutions before evolving into machine learning-based models yet these earlier techniques became inadequate for modern electronic transactions' scale and complexity. Modern credit card fraud detection relies on supervised learning models such as logistic regression, decision trees, support vector machines, and neural networks and also depends on unsupervised and semi-supervised methods when labelled data is scarce. Modern fraud detection techniques rely on deep learning methods including recurrent neural networks and autoencoders because these models discover multiple complex nonlinear patterns and behavioural sequences in transaction data. As financial fraud grows more complex ensemble and hybrid methods receive more usage for improving detection robustness as well as managing multidimensional fraud patterns.

The chapter discussed the ongoing major difficulties which continue affecting both research and practical applications of fraud detection. The effectiveness of predictive modeling suffers from ongoing challenges including class imbalance and label noise as well as limited feature diversity and requirements for real-time decision systems. The research community addressed these problems through the development of resampling methods together with anomaly detection frameworks and cost-sensitive learning approaches. Current research lacks consistent assessment methods for these approaches particularly when used in operational real-world settings. Conventional accuracy measurement tools prove inadequate on unbalanced data distributions so researchers now prefer more informative metrics which include precision and recall alongside F1-score and PR-AUC. The implementation of ML systems at scale shows through PayPal and Mastercard industry examples which integrates performance alongside operational efficiency with regulatory and data protection standards.

The chapter demonstrated technical and operational insights and increasing focus on ethical and legal aspects related to fraud detection systems. The deployment of AI models in financial environments requires organisations to address critical issues which include both algorithmic fairness together with data privacy concerns and explainability standards. Data handling requirements and automated decision justification processes in fraud detection systems are determined by regulatory frameworks which include GDPR, PCI DSS together with national consumer protection laws. Multiple research areas require further attention within the field including unstructured data utilisation, real-time adaptive prevention models, and performance comparisons between different model approaches. The present research seeks to build both theoretical and practical fraud detection advances through robust model comparison of selected ML approaches.

Chapter 3 Methodology

3.1 Chapter Introduction

This chapter describes the methodological framework to explore the issue of the use of ML techniques for detecting credit card fraud. Since fraudulent activities of the digital activities got more complex, there is increasing demand for intelligent, data-driven models that are able to find suspicious patterns with both precision and reliability. The methodology presented in this chapter will facilitate the creation and assessment of various ML algorithms that can be used to tackle the issues of fraud detection in real-world financial systems. The chapter starts by providing the philosophical and the theoretical background for the study, which is of a positivist and quantitative nature to keep it objective and quantifiable. This is followed by a detailed description of what the research strategy might look like and an overview of the kinds of data that are normally seen in fraud detection scenarios.

Some of the pre-processing features that can be applied to various data, including class imbalance, feature scaling, and formatting are discussed, to ready the dataset, for model training. There is also provided a clear training and evaluation plan containing data partitioning, cross-validation and hyperparameter tuning. Lastly, the chapter establishes an approach to analysing the model outputs to produce practicable insights from this exercise. Such a prescriptive pattern ensures systematic and replicable practise of best practises in applied ML for fraudulent detection within the study.

3.2 Research Philosophy

Research philosophy addresses how understanding takes place and which approach is used to undertake research. The chosen philosophy for this frauds detection of credit cards utilising ML is positivism. It is based on the fact the knowledge should be sourced from facts, numbers and data observation and not personal opinions and beliefs (Khanday et al., 2024). Since this study is focused on working with data, modelling and calculating the results based upon the numerical indicators, it is reasonable to prefer positivism as the best fit.

In fraud detection, the aim is to distribute the transactions into two categories: genuine and fraudulent ones by using past data and developing ML algorithms. These decisions are informed by quantitative factors such as the amount and the period of transaction or the frequency of transaction. Evaluations in ML models do not require judgment from humans, as the model will determine the patterns in the data. Their performance is then measured by statistical means such as accuracy, precision, recall and F1-score

(Naidu et al., 2023). This practice is in line with the positivist perspective because all is being measured, tested and repeated in a systematic manner.

Positivism also emphasizes results which can be repeated or generalised. In this research, the purpose is to build a fraud detection solution that can be applicable to various datasets or environments. Through the use of structured method and absence of personal bias, the study is ensured that the findings are creditable and consistent. Therefore, positivism will fit well with the purpose of this research as it will guide the process of modeling and the evaluation of their performance followed by the observations based on the real data.

3.3 Research Strategy

In this study, it will use a deductive strategy of research where it will start with existing theories and models and then apply it through structured approaches. When it comes to credit card fraud detection, numerous studies have evidenced that, using the formerly processed transactions, ML algorithms can be trained to approach fraudulent transactions by merely learning from past records of customers' transactions. Based on this knowledge, the research tries to use particular ML models, for example, logistic regression, random forest, and XGBoost, and determine the extent to which they perform in classifying the transactions. The strategy is valuable in that it enables the researcher to level already-known ideas in a new, broad manner using actual data and performance metrics. Instead of starting with the personal views or observations, the method proceeds from theory to evidence.

The strategy is also useful in the sense that it covers a defined and logical path – from a research problem to an application of proven models to verification of results on the basis of statistical results. It maintains the relevance of the study by ensuring it is concerned with answering a particular question of which ML models are most effective in detecting credit card fraud. A deductive approach also works well with the positivist philosophy and the application of a quantitative method because it is oriented to objectivity and measurement (Pandey, 2019). The models and outcomes can be replicated, checkable and comparable. This means that the research would not only be able to test what has already been proposed in previous literature, but also potentially uncover new findings concerning performance of different techniques. This approach makes deductive strategy facilitate a systemized and credible investigation.

3.4 Data Collection

In research regarding credit card fraud detection with the help of ML, the phase of data collection is an important part. Since the recognition of the patterns requires the financial transactions, the data set has to be structured, numeric, and labelled. In most cases, datasets applied in such studies are obtained from

financial institutions, payment systems, Kaggle, or from the public domain (Babu & Pratap, 2020). Some of such datasets involve transactions' value, transaction time and frequency, as well as anonymized user behaviour features. During collection, the aim is to collect data that has fraudulent and valid transactions for effective training of models. Researchers make sure the dataset has sufficient instances of fraud that can be analysed, although fraud cases are usually scanty.

Since financial data is sensitive in nature, direct access to live or personal credit card transaction records is limited. For this reason, researchers tend to refer to anonymised or synthetic datasets released by secure sources or open ones (Khaled et al., 2024). These datasets are usually cleaned from personal information in compliance with the GDPR data protection laws. Related to working with live data from banks or organizations, data collection typically comes with strict permissions, data use contracts, and undergoing ethical approval. An anonymization technique like tokenization, hashing, or obfuscation of customer information for the purpose of privacy and compliance might also be adopted in the process (Vagadia, 2020).

After being provided with the access to an appropriate data set, the process of data collection will require the confirmation of the structure, completeness and relevance of the information. Considerations include the quality of labelling (fraud and non-fraud) or missing values and class balancing. Researchers also ensure that duplicate records or inconsistencies are checked to affect performance in the model. At this point, data may be brought into analysis tools for further inspection and editing. The procedure of data collection, either through a public source or cooperation with an institution, becomes the foundation of creating reliable and accurate ML models for fraud identification.

3.5 Data Pre-processing Plan

Data pre-processing is a key phase for building up any ML model, in reference to credit card fraud detection, due to the fact that data quality and structure play a significant role in determining how it converts (Alfaiz & Fati, 2022). Raw transaction data sets tend to have imbalances, inconsistencies, and irrelevant or skew data that may destroy the learning process. Not only does the proper pre-processing help to make the model perform better, but it's also necessary to make the outcomes reliable and generalizable.

3.5.1 Data Cleaning and Formatting

The first stage of the pre-processing workflow is cleaning and formatting data (Ahammad et al., 2020). This involves running through the dataset looking for missing value, duplicate records, irrelevant entries,

and inconsistency in the formatting. Small errors in the financial dataset may mislead the results. Missed-out values if any, shall be filled according to context and volume. For the numerical columns that occasionally have missing data, (mean or median) imputation will be applied. However, if some rows or columns consist of too much missing information and are not bringing any meaningful value for fraud detection they can be included as a whole. The transaction entries duplication is also common, and the model can be biased by the pattern which is not a characteristic model.

Moreover, if it is a categorical data, such as transaction type or location, then every such feature must be converted to the numerical value for ML algorithms. This conversion will be achieved by label encoding or one-hot encoding in case the categories are ordinal or nominal respectively (Voican, 2021). Standardizing data types and formats of columns ensures ML libraries compatibility and prepares the ground for pipeline-based automation in further stages.

3.5.2 Feature Scaling and Selection

After cleaning and formatting the data, feature scaling is a mandatory step (Nuthalapati, 2023). This is important for models sensitive to the magnitude of numerical inputs as logistic regression, support vector machine and neural networks. Such attributes as amount of transaction or period can differ by orders of magnitudes in contrast with other attributes, which can bring too many weights to the larger parts of the data, making the model depend too much on large-value features. To avoid this, Min-Max normalization or Z-score standardization will be used, which will place all numerical inputs on the same scale (Jůzová, 2024).

Another important factor is feature selection. Not every feature in a dataset has the same influence on the fraud detection process. Irrelevant or weakly correlated features are capable of adding noise, increasing computational complexity, and causing overfitting. Some of the techniques that will be applied for the identification and retention of the most informative variables include techniques such as: correlation heatmaps, recursive feature elimination (RFE), or using model-based importance scores (for instance, using the random forest or XGBoost feature rankings) (Priscilla & Prabha, 2021). This allows simplification of training and aids in letting models understand the most important fraud signals to act upon.

3.5.3 Handling Class Imbalance

A considerable problem in fraud detection is that almost every dataset exhibits severe class imbalance. Fraudulent transactions typically form less than 1% of the total records, which makes it hard for models to

detect them without specific intervention. If this imbalance is not corrected, a model could predict almost all transactions as non-fraudulent and still have high accuracy, but at the cost of failing completely its purpose.

To address it, resampling techniques will be used during pre-processing. One of the popular methods is SMOTE, it generates new synthetic fraud cases by interpolating between existing minority samples (Parkinson de Castro, 2020). ADASYN utilizes the idea of SMOTE further with the goal of learning from the examples that are hard to learn. Under-sampling the majority class is also possible if such a dataset is large enough to spare a part of it without significant loss. The application of these techniques will make the model be subjected to a combination of fraud and non-fraud cases, which will enable it to identify rare but important events.

3.5.4 Pipeline Automation for Reproducibility

To ascertain all the pre-processing steps are effectively applied in the process, a pipeline approach will be used. This encompasses the ability to integrate the pre-processing operations like scaling, encoding, resampling, and feature selections into a processing pipeline for reuse of the whole sequential process. Pre-processing techniques are easily applicable using methods from the Scikit-learn pipeline and imblearn.pipeline.Pipeline, each step can be run in a fixed sequence and iteration occurs when the data is reset or if a new model is trained (Ahmed & Green II2022).

Automated pipelines offer multiple benefits, they minimize the intervention of human values, make it easy to reproduce results whenever the models or datasets are changed (Dickerson & Worthen, 2024). For example, if a new dataset of fraud is used, then it is convenient to repeat all the operations without having to redo it step by step. This also enhances the level of transparency as every operation that is performed on the data is stated and recorded in the pipeline format. In future deployment, pipeline automation would be important in expanding the fraud detection system and also incorporating with other tools that deal with real time transaction monitoring.

3.6 ML Models

3.6.1 Proposed Algorithms

This research aims to detect Credit Card fraud detection using ML algorithms including Logistic Regression, Random Forest, XGBoost, and ANNs. These models were used because they are widely used in the literature and found to be suitable to deal with various data characteristics. Logistic Regression is one of the commonly used statistical model for regression analysis which is often used for binary

classification method (Alenzi & Aljehane, 2020). It estimates the conditional probability of a class label as a function of the input variables through the log-odds of the outcome. Despite its basic model structure, it is used in fraud detection due to its advantage of high interpretability and easy implementation. Random Forest is a type of method which builds a multiple tree and puts together their results to achieve a more stable decision. It is good for non-linear data and can provide outputs on the significance of each feature for further prediction and analysis of the model.

XGBoost is an advanced ensemble of the decision-tree-based model, where boosting is conducted in a sequential way (Mohbey et al., 2022). Every new tree tries to fix the mistakes of the previous one and the model can learn over existing errors. It incorporates applications such as regularization techniques that minimize cases of overfitting and is known to run efficiently in terms of technique performance thereby widely used in fraud detection studies. ANN based on the structure of the human brain can model complex, non-linear relationships by channelling data through various hiding layers of interconnected neurons (Poddar, 2024). ANNs are able to pick up on very subtle trends in transaction behaviour that simpler models may lack. However, they do need much data, more learning time, and careful tuning of parameters. These models provide a balanced combination of interpretability, robustness, and complexity. Therefore, their presence provides the option for an inclusive comparison of models across different learning efforts and the strengths of algorithms.

3.6.2 Model Training Plan

That a structured model training plan will be followed to guarantee that evaluation of each ML algorithm is done fairly and uniformly. The data set will also be split into training and test subsets in a standard split ratio, for example, 70% of the data set for training, 30% for testing (Genc et al., 2019). Such a split enables the models to learn from most of the data and set aside part of it for independent evaluation. Sampling can be utilised for the split in order to preserve the initial distribution of fraudulent and non-fraudulent classes, as fraud datasets are usually highly imbalanced (Muaz et al., 2020). This avoids biasedness of model assessment and the testing set is made similar to real condition. Furthermore, cross-valuation will be used to establish the stability of the models' performance. A fold (5-fold or 10-fold) will be used in order to split the training data to smaller sets for fitting of the model in each of these folds and validated against. This leads to over-fitting resolution and improved estimation of performance of the model in unseen data.

To take the performance even further, hyperparameter tuning will be applied on each of the models. This is identifying the best input for some of the critical parameters that influence the model's behaviour. For

example, on Random Forest, a parameter tuning by raising trees, maximum depth and minimum samples per leaf will be tuned. Tuning in XGBoost will be carried out for various parameters including learning rate, depth of the tree and subsample ratios (Probst et al., 2019). The neural networks will be optimised by tuning of their hidden layer, neurons per layer, activation functions and learning rates. The tuning process is achieved by the utilisation of either GridSearchCV or RandomizedSearchCV that are used to test various combinations of parameters by way of cross-validation to achieve the configuration that achieves the best results (Arshad et al., 2024). Through periodic training and such working of tuning, every model will get a chance to achieve its best performance, and their performance result can be compared on equal footing.

3.7 Data Analysis

Performance evaluation on the basis of prediction of the ML models as well as discovering insights are very important, which occur on the stage of the data analysis. Not only it provides statistical validation of the models, but also interpretation value capable of justifying real world decision making.

3.7.1 Evaluation Metrics and Performance Assessment

The effectiveness of each model will be determined under a set of evaluation metrics that will give insight on the extent to which fraudulent and non-fraudulent transactions are classified. These consist of accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC (Btoush et al., 2023). Even though accuracy is the most often used metric, it can be misleading in a highly unbalanced dataset, which often occurs with fraud detection. Consequently, precision and recall will be prioritized. Precision is the rate of the correctly identified frauds relative to all predicted frauds, while recall is the rate of correctly identified actual frauds. The F1-score takes both metrics and provides an equilibrated perspective of performance especially when the cost of false negatives and false positives are both high.

Moreover, ROC-AUC and PR-AUC will be used in order to assess the discrimination power of the models. ROC-AUC shows how well model separates classes at different threshold, whereas PR-AUC is helpful for imbalanced datasets, indicating performance on the minority class (Richardson et al., 2023). These metrics will determine which models are upholding high fraud deterring capability without being burdened by non-fraud cases. The most effective model in this study will be considered the one based on the best combination of F1-score and PR-AUC.

3.7.2 Confusion Matrix and Error Analysis

In order to understand the errors committed by each of the models, confusion matrix will be created and also analysed. This matrix gives the values of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) (Obi, 2023). In the scope of detecting frauds, false positive means the authenticity of the transaction which is flagged as a fraud, while false negative means a fraudulent transaction is missed. Although the cost of a false positive and a false negative has real world implications, the cost of a false negative is generally higher because then lost fraud and possible financial loss occurs.

Analysis of the confusion matrix makes it easier to understand the strengths and weaknesses of every model (Amin & Mahmoud, 2022). A model that has high recall but low precision, for instance can be flagged multiple genuine transactions as being fraudulent hence it would need to be manually reviewed. On the other hand, even a model that has low precision but high recall may not pick up fraudulent activity at all. In connection to the above metrics, the confusion matrix will add a full detail on how each model will react to realistic conditions. Some of these visualizations, such as heatmaps, will be used to display these matrices in easy-to-read formats for comparative analysis.

3.7.3 Feature Importance and Model Interpretability

It is also important to understand why a model makes certain predictions apart from performance metrics. Feature importance analysis will be performed especially for tree-based models such as Random forest and XGBoost which will enable the quantifying of contribution of individual input variable (Zhou & Hooker, 2021). This will allow identifying which variables like transaction amount, time or frequency are the most effective in forecasting fraud trends. All these factors make the model transparent and the outputs more credible.

In more advanced models such as neural networks, which are often called “black boxes,” other tools for interpretability are available as well (Agarwal & Ratha, 2021). SHAP (SHapley Additive exPlanations) values can be used to show the influence of individual features on the model’s output for given transactions. Not only do these insights facilitate validation but they help to identify any unusual behaviours or patterns that can be applied to prevent fraud in a better way. The interpretability of model supports informed decision-making where analysts as well as stakeholders can put faith and act on the predictions.

3.7.4 Translating Results into Business Insights

The end goal of fraud detection models is not for prediction alone and rather for the purpose of guiding business decisions. Upon assessment of model performance and interpretation of key features, the

findings will be translated into executable insights. For example, the business may introduce location-based verification protocols for cases, where there are constantly out-of-area transactions flagged as fraud. Likewise, when models indicate a high occurrence of fraud at certain times, fraud monitoring systems could then be heightened during those times.

In addition, the analysis of false positives and false negatives can help businesses to understand where to spend money, for instance, in increased customer verification processes, employee training, or alert systems. This kind of analysis also promotes policy development on threshold setting, the determination of what to block on automatically and when to pass it through with a flag. It aims to minimize financial loss, enhance customer experience and ensure that fraud detection does not just give results but also fit into the business and risk tolerance.

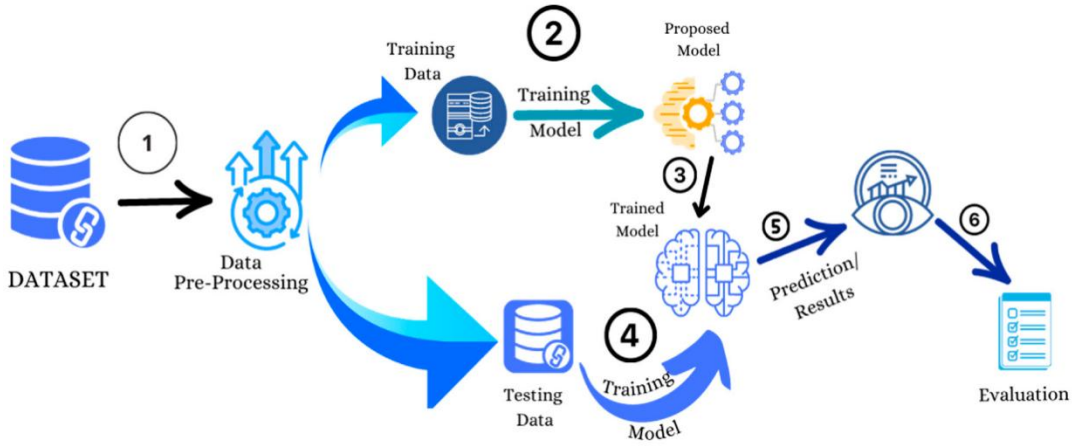

Figure 8: Overall ML workflow (Khalid et al., 2024)

3.8 Chapter Summary

This chapter provided the methodology that used to facilitate the detection of credit card fraud by machine learning. It started by developing research philosophy and approach which is based on positivism and quantitative framework. A deductive approach was chosen, based on using ML algorithms that already exist to apply such algorithms to structured transaction data. Despite the lack of a definite dataset, the chapter provided the general properties of data for fraud detection, as numerical features and class imbalance. These attributes guided the pre-processing and modeling choices. The methodology was organized in such a way that the process can be objective, measurable, and reproducible.

An overall data pre-processing plan was outlined with the steps for cleaning formatting, and preparing the data for training models. This involved dealing with missing data, scaling the features to standard range, and choosing the most relevant variables for classification. Particular focus was put on treating class

imbalance with resampling approaches such as SMOTE and ADASYN to give the minority class (fraud) representation. Moreover, the incorporation of the pipeline automation became a way to simplify the pre-processing process and secure consistency within the model training runs. These steps were done to minimize the manual errors and enhance reproducibility in the experimental phase.

The ML models that were chosen for evaluation included Logistic Regression, Random Forest, XGBoost, and Artificial Neural Networks. Such models provide a combination of simplicity, interpretability, and complexity to allow for a balanced comparison. Data splitting with cross-validation and hyperparameter tuning with GridSearchCV were included in the training plan. Model outputs analysis ways such as confusion matrix interpretation and feature importance were among the last parts of the chapter. The analysis plan intended to consider how the result would be transformed into business-relevant insights in order to ensure the compliance of the model performance with the requirements of fraud detection in the real world.

Chapter 4 Results and Discussion

4.1 Chapter Introduction

This chapter presents the implementation results of the study that are focused on the development and testing of ML models to identify fraud cases in credit card data. The chapter details the steps of the modelling process including data preparation, exploratory analysis, and the problem about class imbalance and feature scaling. Such steps were necessary in order to maintain the integrity and reliability of data used in all experiments.

A variety of classification models had been applied with a view to assessing their performance in identifying fraud in the case of the imbalanced data set. Performance of the model was initially analysed using common metrics and later corrections were made to enhance precision of predictions. Different visualisations such as confusion matrices, ROC curves, and precision-recall plots were used to allow better visual understanding of the results of investigation. In addition, the analysis of feature importance identified the most important variables strongly affecting fraud detection. Finally, all different models were compared in-depth to figure out which of them performs best in terms of performance metrics and the actual practises in the real world.

4.2 Data Preparation & Pre-processing

Data preparation must be effectively handled so that successful outcomes are attained by any ML pipeline with a high frequency of imbalanced datasets which characterises fraud detection applications (Kaur et al., 2019). The pre-processing and initial preparation efforts carried out prior to the model development are shown as follows. It includes data set structure understanding, data checking, imbalanced data treating, split the data set correctly and feature normalisation. Through systematic implementation, the procedures were implemented such that all models processed balanced and clean data equally, enhancing performance and precise learning.

4.2.1 Dataset Description

The dataset has been chosen from Kaggle, it consists of anonymized credit card transaction data (DATASET). This dataset is useful for understanding methods of detecting fraudulent activities, and this makes it a critical resource for producing and testing predictive models in fraud detection. This work used a real-world credit card transaction data set consisting of 31 variables and 284,807 transactions. The dataset has 28 principal components achievable through PCA, V1 to V28, with features such as, ‘Time’, ‘Amount’, and target variable ‘Class’. The ‘Class’ variable is binary, with 0 indicating a legitimate

transaction and 1 indicating a fraudulent one. Due to confidentiality concerns, the original features have been anonymized. The dataset is heavily imbalanced, with only 492 fraudulent transactions, representing approximately 0.172% of all data. This class imbalance poses a significant challenge to ML models, as they may become biased toward predicting the majority class. The dataset was loaded into the Python environment using the pandas library and was initially examined for completeness and consistency.

Table 1: Data types and non-null counts in the dataset

#	Column	Non-Null Count	Dtype
0	Time	284807	float64
1	V1	284807	float64
2	V2	284807	float64
3	V3	284807	float64
4	V4	284807	float64
5	V5	284807	float64
6	V6	284807	float64
7	V7	284807	float64
8	V8	284807	float64
9	V9	284807	float64
10	V10	284807	float64
11	V11	284807	float64
12	V12	284807	float64
13	V13	284807	float64
14	V14	284807	float64
15	V15	284807	float64
16	V16	284807	float64

17	V17	284807	float64
18	V18	284807	float64
19	V19	284807	float64
20	V20	284807	float64
21	V21	284807	float64
22	V22	284807	float64
23	V23	284807	float64
24	V24	284807	float64
25	V25	284807	float64
26	V26	284807	float64
27	V27	284807	float64
28	V28	284807	float64
29	Amount	284807	float64
30	Class	284807	int64

Table 2: descriptive statistics of dataset.

Statistic	**Time**	**V1**	**V2**	**V3**	**V4**	**V5**	**...**	**V28**	**Amount**	**Class**
count	284807	284807	284807	284807	284807	284807	...	284807	284807	284807
mean	94813.86	-1.101	0.168	0.045	-0.121	0.066	...	0.000	88.35	0.0017
std	47488.15	1.960	1.651	1.516	1.415	1.380	...	0.330	250.12	0.037
min	0.0	-56.407	-72.715	-48.325	-5.6831	-113.74	...	-15.430	0.00	0.0

		5	7	6		3		1		
25%	54201.0	-0.9204	-0.5985	-0.8904	-0.8486	-0.6916	...	-0.3973	5.60	0.0
50%	84692.0	0.0181	0.066	-0.0321	0.0181	-0.0297	...	0.0529	22.00	0.0
75%	139320.0	1.3156	0.8037	1.0272	0.7439	0.6069	...	0.4799	77.17	0.0
max	172792.0	2.4549	22.0577	9.3826	16.8753	34.8017	...	3.3847	25691.16	1.0

Table 3: Class imbalance before SMOTE application

Class	Count
0	284,315
1	492

4.2.2 Missing Value Check and Data Types

A necessary first step in data preparation was validating for missing/null values (Khare & Srivastava, 2023). The dataset was found to not have any missing values thus no imputation or deletion was needed. Moreover, all feature columns were numerical, so no encoding was needed. At this initial stage, by guaranteeing data integrity, it was possible to reduce the possibility of making errors in the model training stage.

images:

Table 4: Missing value check across all features.

Column	Missing Values
Time	0
V1	0

V2	0
V3	0
V4	0
V5	0
V6	0
V7	0
V8	0
V9	0
V10	0
V11	0
V12	0
V13	0
V14	0
V15	0
V16	0
V17	0
V18	0
V19	0
V20	0
V21	0
V22	0
V23	0
V24	0

V25	0
V26	0
V27	0
V28	0
Amount	0
Class	0

4.2.3 Addressing Class Imbalance

Due to the extreme degree of imbalance of classes, SMOTE was employed to balance the training set. SMOTE operates by creating synthetic cases of the minority class using feature-space similarities (Mukherjee & Khushi, 2021). This allows models to learn the decision boundary of rare events such as fraud without just duplicating records. Before using SMOTE, the stratified 70:30 train-test split was used to maintain the original dataset's data distribution of classes. SMOTE was then used only on the training data to avoid data leak.

Figure 13: Class balance before applying SMOTE. (bar chart).

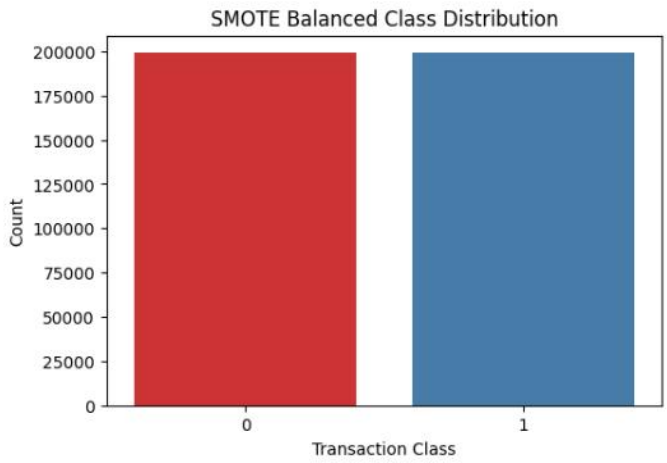

Figure 14: Class balance after applying SMOTE.

4.2.4 Train-Test Split

The dataset was divided using scikit-learn's train_test_split() function with a stratified split. Stratification ensures that the proportion of fraud and non-fraud transactions is maintained in both training and test datasets. This is especially important for imbalanced datasets to ensure the test set reflects the real-world distribution (Tan et al., 2021). After the split, 70% of the data was used for training and 30% for testing. The training data was used for SMOTE, scaling, and model training, while the untouched test data was used exclusively for evaluating model performance.

```
X = df.drop('Class', axis=1)
y = df['Class']
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42, stratify=y)

smote = SMOTE(random_state=42)
X_train_res, y_train_res = smote.fit_resample(X_train, y_train)

print("Original training class distribution:")
print(y_train.value_counts(), "\n")

print("Resampled training class distribution (after SMOTE):")
print(pd.Series(y_train_res).value_counts())
```

```
Original training class distribution:
Class
0    199020
1       344
Name: count, dtype: int64

Resampled training class distribution (after SMOTE):
Class
0    199020
1    199020
Name: count, dtype: int64
```

Figure 15: SMOTE Resampling and Class Distribution Output.

4.2.5 Feature Scaling

Due to the wide range of values in features like ‘Amount’ and ‘Time’, standardization was performed using StandardScaler. This step was critical for models such as Logistic Regression and ANN, which are sensitive to feature magnitude. The scaler was fit only on the training data to prevent data leakage and then applied to both training and test sets.

Before Scaling

Index	Time	V1	V2
0	154640.0	-0.012102	0.707332
1	139525.0	1.776151	-0.184642
2	69778.0	-1.083391	-4.440527
3	48473.0	-0.518847	1.025087
4	129350.0	-0.640421	0.212171

After Scaling

Index	Time	V1	V2
0	1.383197	0.449824	-0.343830
1	1.068078	0.775656	-0.591649
2	-0.386012	0.254627	-1.774070
3	-0.830181	0.357491	-0.255547
4	0.855949	0.335340	-0.481401

Figure 16: Feature Scaling output.

4.3 Exploratory Data Analysis (EDA)

It is important first to understand the data’s underlying structure and patterns before using it into the ML models. EDA was performed to determine key trends, examine how features correlate, detect outliers, and

investigate class behaviour (Patel & Patel, 2024). Given that the dataset is heavily imbalanced and anonymized in nature, visualization takes on a central role in identifying latent trends, and in suggesting choices of the relevant features and modeling approaches.

4.3.1 Distribution of Transaction Amounts

The distribution of the Amount feature in the dataset shows a very skewed distribution. A boxplot was used to visualize the skewness and outliers. Most of the legitimate transactions are within the relatively low amount range and fraudulent transactions are scattered and include several high-value anomalies. This biased distribution led to subsequent use of feature scaling and justified the consideration of Amount as a potential important variable in fraud detection.

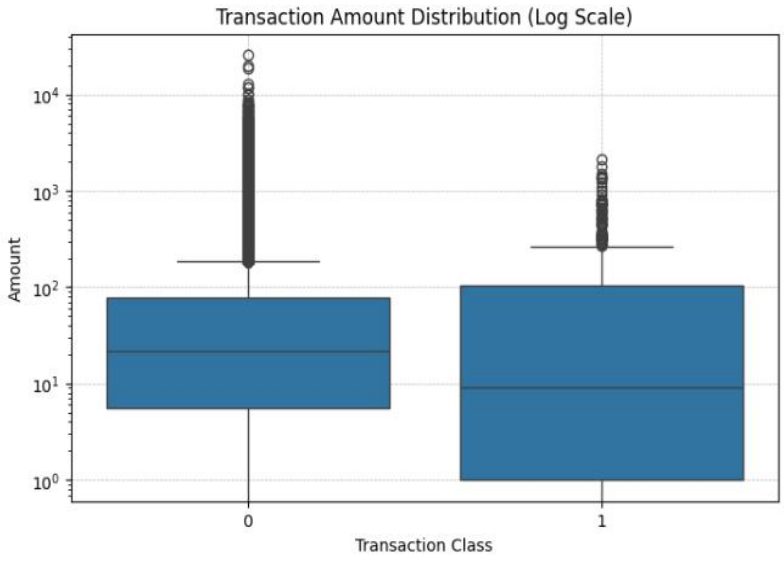

Figure 17: Boxplot of transaction amounts by class.

4.3.2 Class-wise Distribution (KDE Plot)

To understand how transaction patterns differ between fraudulent and legitimate transactions, Kernel Density Estimation (KDE) plots were generated. These plots showed stark differences in the distribution of specific principal components (e.g., V14, V10, V17) between Class 0 and Class 1. Such visual cues were early indicators that certain features carry meaningful discriminatory power and were later validated during model training.

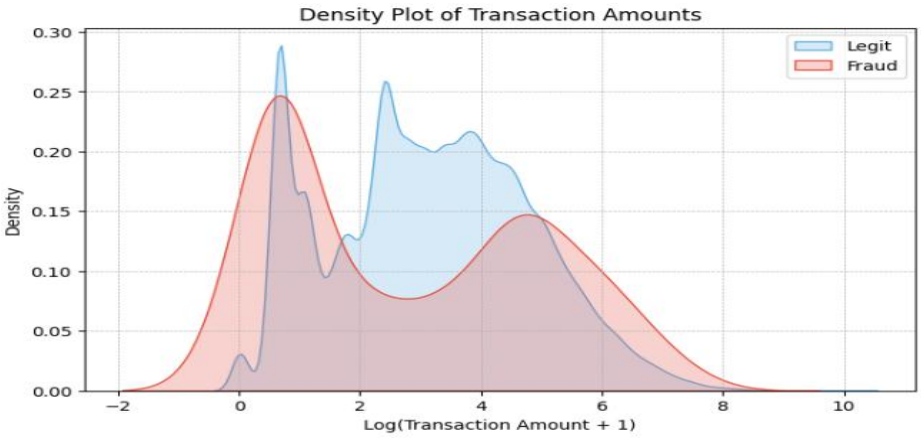

Figure 18: KDE plot of transaction amounts.

4.3.3 Correlation Analysis

A correlation heatmap was generated to observe relationships among features. Since the dataset is mostly comprised of PCA-transformed features, the majority of correlations are weak, as expected. However, the heatmap helped confirm that no two features are overly collinear, reducing the risk of multicollinearity. It also helped reinforce which features may be more independent and, thus, useful for models that rely on orthogonality like Logistic Regression.

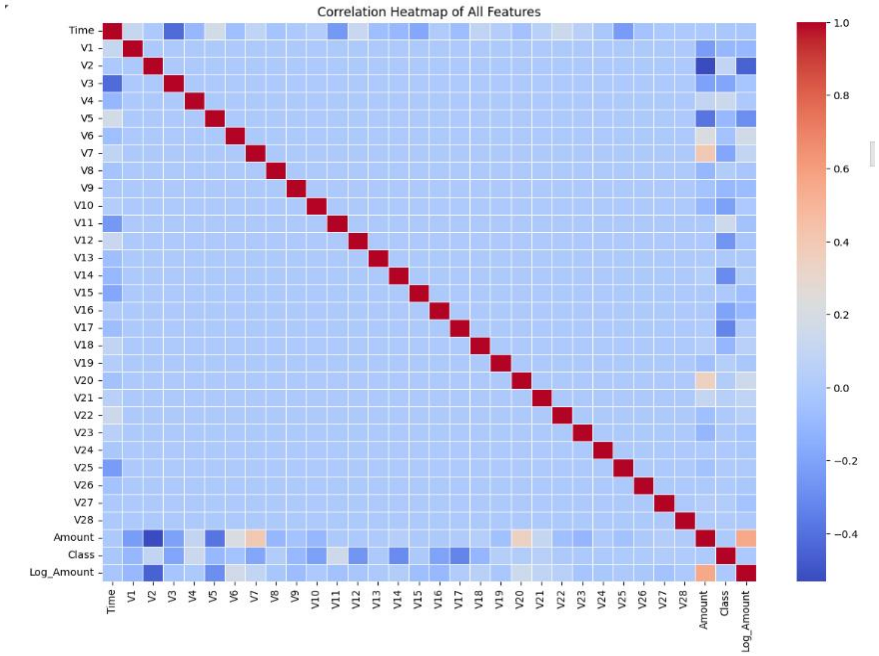

Figure 19: Correlation heatmap of all features.

4.3.4 Fraud Rate by Hour of Transaction

The Time feature, representing seconds elapsed since the first transaction, was converted to "hour" format and visualized to identify fraud trends over time. The fraud rate per hour revealed an interesting spike during certain times of the day. This temporal trend suggests that fraudulent behaviour may be more frequent during specific time windows, potentially correlating with lower monitoring periods or vulnerabilities in banking operations.

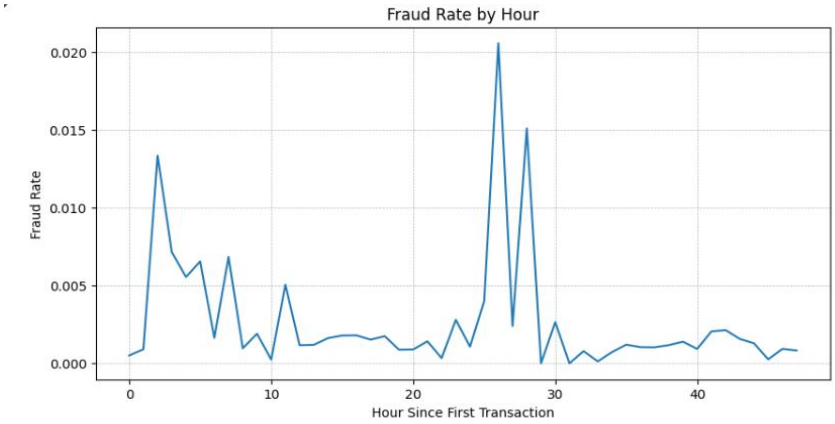

Figure 20: line chart Fraud rate across hours since first transaction.

Exploratory analysis uncovered notable patterns between legitimate and fraudulent transactions. Fraudulent cases often involved lower amounts and displayed distinct density distributions compared to non-fraudulent ones. The correlation heatmap revealed that most features were uncorrelated, supporting the use of tree-based models. Time-based analysis showed certain hours with elevated fraud rates. These insights informed both pre-processing strategies and model interpretation in later stages.

4.4 Baseline Model Implementation

To evaluate the effectiveness of different classification techniques in identifying fraudulent credit card transactions, four baseline models were implemented: Logistic Regression, Random Forest, XGBoost, and ANN. These models were first trained using default hyperparameters to establish baseline performance benchmarks. Each model was evaluated using standard classification metrics including precision, recall, F1-score, and AUC-ROC, with particular attention given to the minority class (fraud). Confusion matrices, ROC curves and precision recall plots were also created to support the quantitative

analysis. The configuration, results, and early findings for each individual model are presented in the following section.

4.4.1 Logistic Regression

Logistic Regression was used as a baseline model on which more sophisticated classifiers could be evaluated. Applying SMOTE for dataset balancing, the model was constructed with a maximal iteration of 1000 and a set predefined random seed for repeatability. StandardScaler was used to scale the features to equalise the impact of input variables on the model's decision boundary. Classification metrics such as precision, recall, F1-score, and overall accuracy were reported in the model after training it to the test set. This provided a critical benchmark, illustrating the difficulty of linear models when presented with nonlinear and imbalanced data with its simple, clear structure.

In terms of performance, Logistic Regression achieved an overall accuracy of 0.9983. However, its effectiveness in detecting fraudulent transactions (Class 1) was notably limited. The precision for Class 1 stood at only 0.1488, while recall reached 0.8514, resulting in a low F1-score of 0.2533. This indicates that while the model managed to catch a relatively high portion of actual fraud cases (recall), it also misclassified many legitimate transactions as fraud, leading to a high false positive rate. The confusion matrix supported this, showing 126 true positives but also 721 false positives—a substantial cost in real-world applications where false alarms burden investigation teams. In contrast, for non-fraudulent transactions (Class 0), the model maintained very high precision (0.9997), recall (0.9915), and an impressive F1-score of 0.9956. The ROC curve for the model yielded an AUC of 0.9665, suggesting the model was capable of separating the two classes reasonably well in terms of probability estimates. However, the precision-recall curve for Class 1 confirmed the performance imbalance. Despite being computationally efficient and easy to deploy, Logistic Regression's linear nature limited its ability to capture the complex relationships inherent in fraudulent behaviours, especially when features were transformed PCA components with nonlinear patterns. Its performance clearly indicated the need for more sophisticated models capable of managing class imbalance and nonlinear separability more effectively.

Table 5: Classification report for logistic regression.

Class	Precision	Recall	F1-Score	Support
0	0.9997	0.9915	0.9956	85,295
1	0.1488	0.8514	0.2533	148

Accuracy			**0.9913**	85,443
Macro Avg	0.5743	0.9214	0.6244	85,443
Weighted Avg	0.9983	0.9913	0.9943	85,443

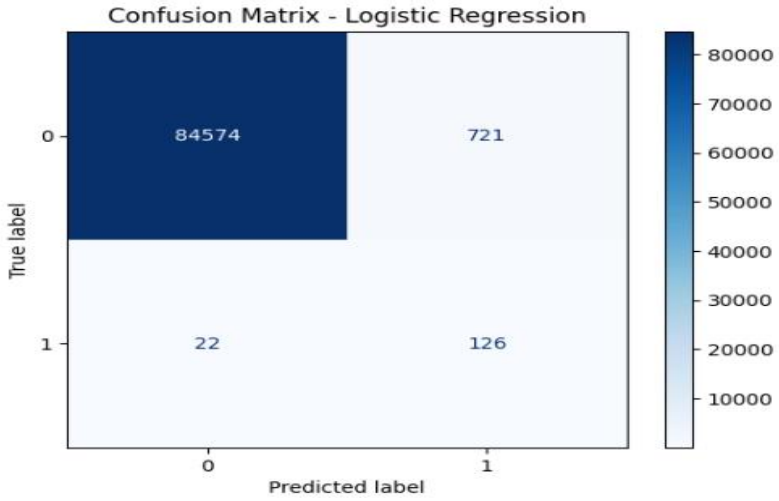

Figure 22: Confusion matrix for Logistic Regression.

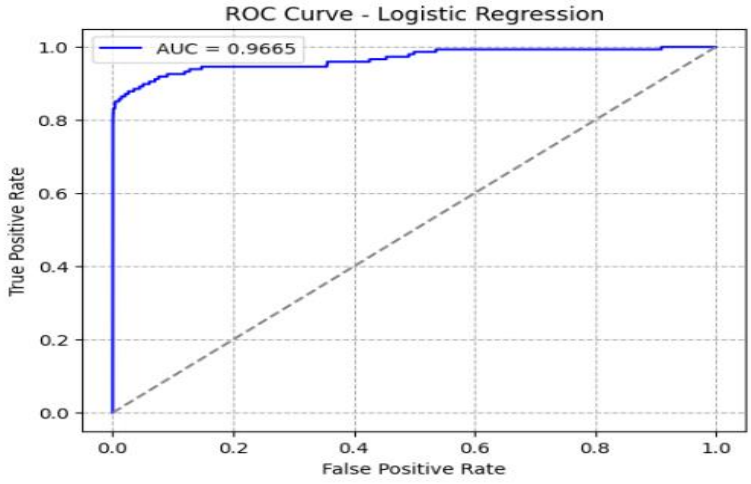

Figure 23: ROC curve – Logistic Regression.

4.4.2 Random Forest

The Random Forest model was implemented as a powerful ensemble method to address the limitations observed in the logistic regression baseline. Utilising 100 decision trees (n_estimators = 100) with a fixed random seed for consistency, the model was trained on the SMOTE-resampled dataset with scaled features. This approach leveraged the strengths of bagging and feature randomness to build a collection of de-correlated trees, thereby improving generalization and reducing the likelihood of overfitting. Random Forest is inherently robust to outliers and multicollinearity, making it particularly suitable for datasets like this one, where the original features (V1–V28) are PCA-transformed. The classifier was evaluated on the test set, and several key performance metrics were recorded, including accuracy, class-specific precision and recall, F1-scores, and visualizations through confusion matrix and ROC/PR curves.

The model delivered excellent results overall, achieving an impressive accuracy of 0.9995. For the non-fraudulent class (Class 0), the precision, recall, and F1-score were exceptionally high at 0.9996, 0.9998, and 0.9997, respectively. More notably, Random Forest made significant improvements in detecting fraudulent transactions (Class 1) compared to logistic regression. It achieved a precision of 0.8806, a recall of 0.7973, and an F1-score of 0.8369, indicating a much more balanced performance between identifying true frauds and minimizing false positives. The confusion matrix shows that the model correctly identified 118 out of 148 fraud cases, with only 30 false negatives and 16 false positives—remarkably low given the size and imbalance of the dataset. The model's ROC AUC score of 0.9493 reflects strong class separability, while the precision-recall curve also confirmed its reliability under class imbalance. Additionally, the feature importance plot highlighted V14, V10, and V4 as the most influential predictors, reinforcing earlier insights from the EDA phase. These features consistently demonstrated predictive power across different models, further validating their relevance. Overall, Random Forest emerged as a highly capable model with strong fraud detection capabilities, good generalization, and valuable interpretability through feature rankings—making it a suitable candidate for practical deployment in real-world fraud detection systems.

Table 6: Classification report Random Forest.

Class	Precision	Recall	F1-Score	Support
0	0.9996	0.9998	0.9997	85,295
1	0.8806	0.7973	0.8369	148
Accuracy			**0.9995**	85,443
Macro Avg	0.9401	0.8986	0.9183	85,443

Weighted Avg	0.9994	0.9995	0.9994	85,443

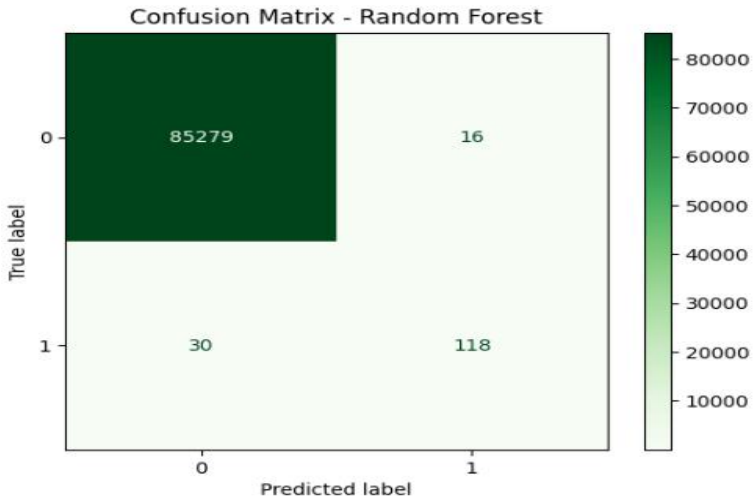

Figure 25: confusion matrix for Random Forest.

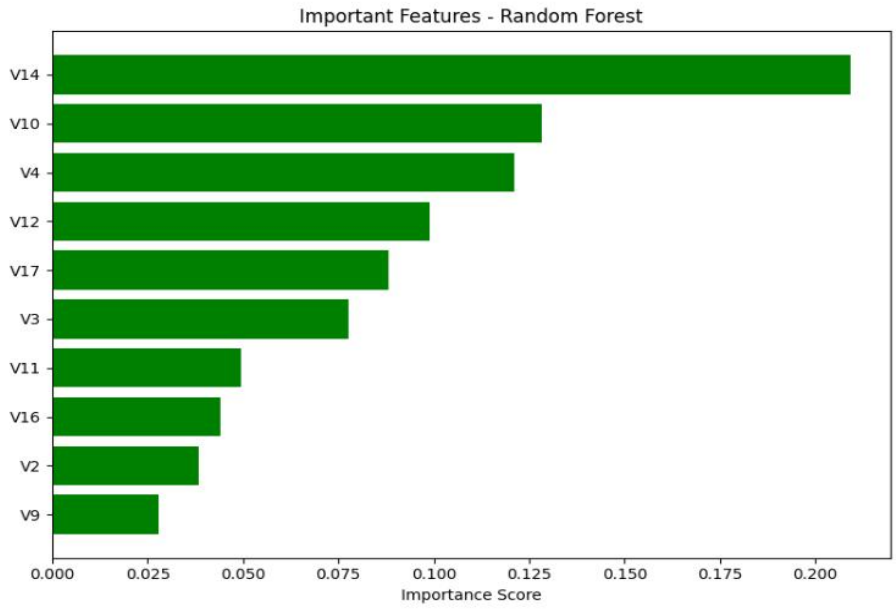

Figure 26: Bar chart Feature importance - Random Forest.

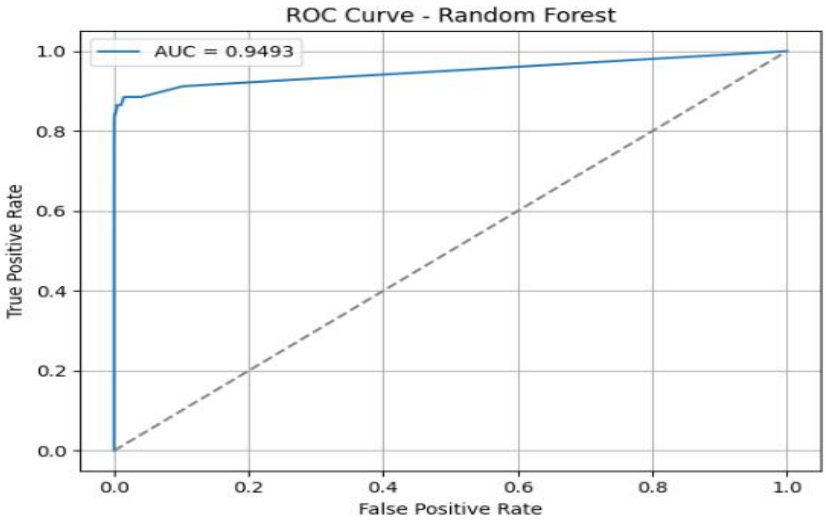

Figure 27: ROC curve – Random Forest.

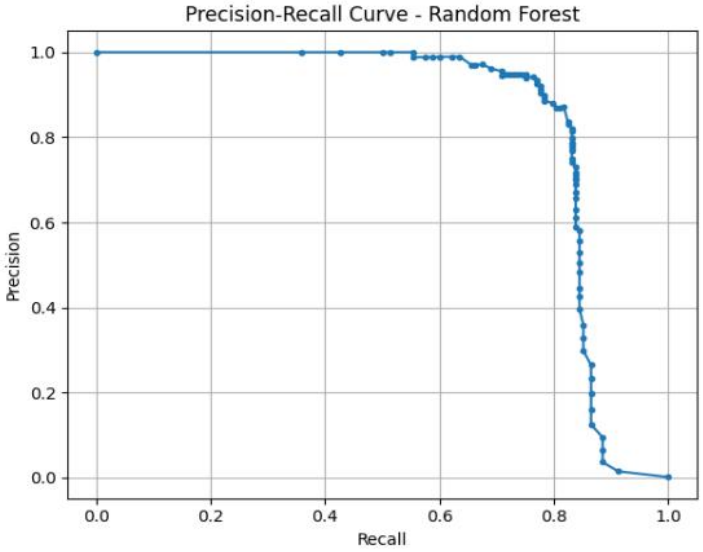

Figure 28: Precision-Recall curve – Random Forest.

4.4.3 XGBoost

The XGBoost model was selected for its reputation as one of the most efficient and high-performing gradients boosting algorithms, especially in structured tabular data tasks like fraud detection. The model was configured with key hyperparameters such as n_estimators = 200, max_depth = 5, learning_rate = 0.1, and subsample = 0.8, optimizing both depth and breadth of learning while controlling for overfitting. The

training was performed on the 50,000-sample SMOTE-resampled dataset with scaled features, ensuring computational feasibility while preserving class balance. Unlike Random Forest, which builds trees in parallel, XGBoost constructs trees sequentially, each correcting the residuals of its predecessor. This allowed XGBoost to fine-tune its learning on the misclassified points, particularly beneficial in detecting minority class instances like fraud. Evaluation metrics included accuracy, precision, recall, F1-score, ROC AUC, PR curves, confusion matrix, and feature importance.

The model demonstrated exceptional performance, achieving an overall accuracy of 0.9994, which is consistent with other top-performing classifiers in this study. For the fraud class (Class 1), XGBoost recorded a precision of 0.8429, recall of 0.7973, and an F1-score of 0.8194. These values indicate that the model strikes a solid balance between correctly identifying fraudulent transactions and limiting false positives. It detected 118 out of 148 fraudulent cases with 30 false negatives and 22 false positives, reflecting strong detection capabilities in a highly imbalanced setting. The AUC score of 0.9779 further emphasizes its excellent discriminatory ability. The PR curve maintained high precision even at moderate recall levels, underscoring its effectiveness in high-risk financial applications where minimizing false positives is as critical as detecting fraud. The feature importance plot confirmed V14 as the most dominant predictor by a large margin, followed by V4, V12, and V17, aligning with both domain intuition and findings from Random Forest. The sharp concentration of feature importance in V14 suggests that certain latent features derived from PCA transformations carry substantial predictive power. Overall, XGBoost's robust performance, particularly in handling imbalanced data with high precision and recall, reinforces its position as one of the most viable models for real-time fraud detection systems.

Table 7: Classification report for XGBoost

Class	Precision	Recall	F1-Score	Support
0	0.9996	0.9997	0.9997	85,295
1	0.8429	0.7973	0.8194	148
Accuracy			**0.9994**	85,443
Macro Avg	0.9213	0.8985	0.9096	85,443
Weighted Avg	0.9994	0.9994	0.9994	85,443

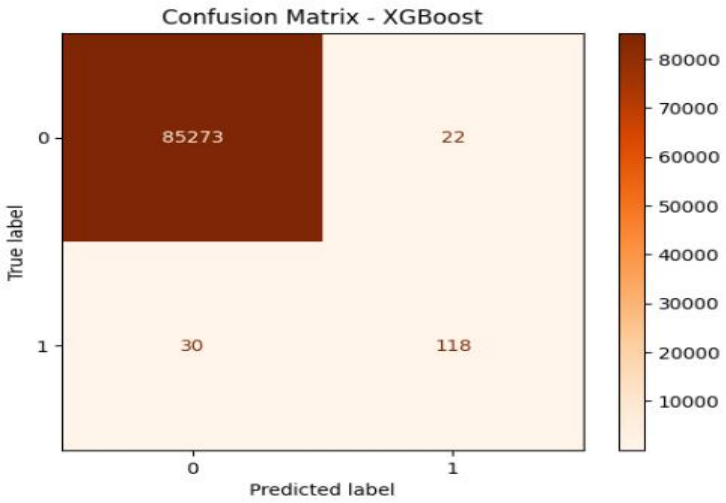

Figure 30: confusion matrix – XGBoost.

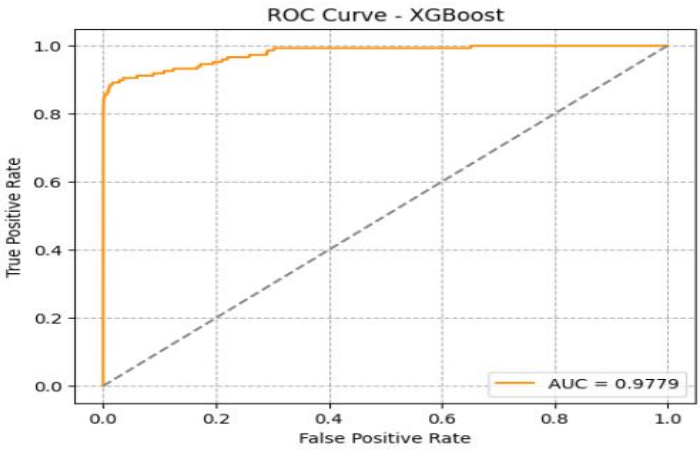

Figure 31: ROC curve – XGBoost.

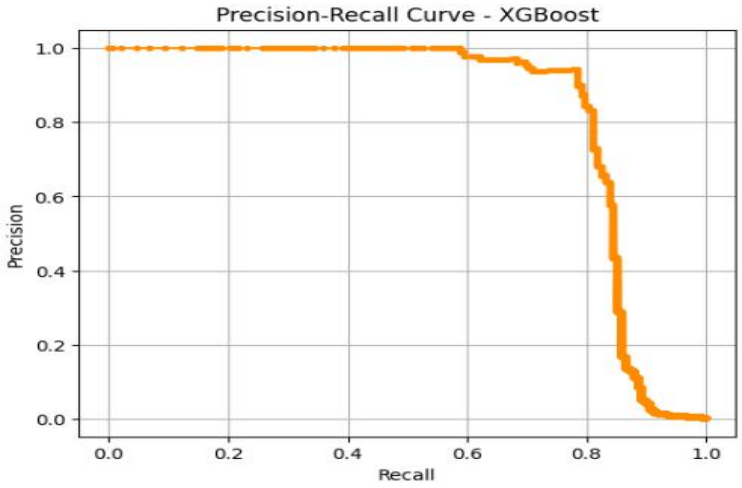

Figure 32: Precision-Recall curve – XGBoost.

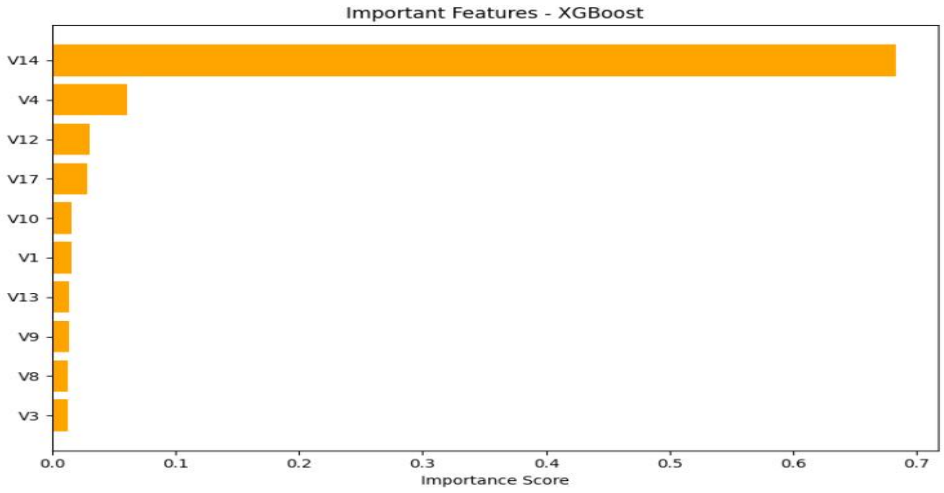

Figure 33: Feature importance – XGBoost.

4.4.4 Artificial Neural Network (ANN)

ANNs were implemented due to their capability to model complex, nonlinear relationships in data. The architecture was constructed using a sequential model comprising multiple dense layers. Specifically, the model included an input layer followed by two hidden layers, each using the ReLU activation function, and a final output layer with a sigmoid activation to produce a binary classification probability. Dropout layers were integrated between the dense layers to prevent overfitting and promote generalization. The model was compiled using the binary cross-entropy loss function and optimized using the Adam

optimizer. It was trained on the SMOTE-resampled dataset consisting of 50,000 instances, with a batch size of 256 and for 12 epochs. Feature scaling using StandardScaler was crucial before training, as ANNs are sensitive to the scale of input features.

Upon evaluation on the original test set, the ANN achieved an overall accuracy of 0.9988, which is consistent with the high accuracy achieved by other models. For the minority class (fraud), the ANN achieved a precision of 0.6269, a recall of 0.8176, and an F1-score of 0.7097. Although the model detected 121 out of 148 fraudulent transactions, it also flagged 72 legitimate transactions as fraud, leading to a relatively higher false positive rate compared to Random Forest and XGBoost. This is evident in the confusion matrix, which showed 27 false negatives and 72 false positives. The ROC curve for the ANN indicated strong separability between the two classes, with an AUC of 0.9642, while the precision-recall curve illustrated a moderate balance, confirming the model's ability to maintain recall but at the cost of some precision. Although ANN performance fell slightly behind Random Forest and XGBoost in terms of fraud-specific F1-score, it still outperformed Logistic Regression by a wide margin. The model's stability across epochs and resilience to overfitting, aided by dropout regularization, confirmed its suitability for the task. However, the lack of interpretability and higher computational demands compared to tree-based models are considerations that must be considered when deploying such models in real-time or high-stakes financial environments.

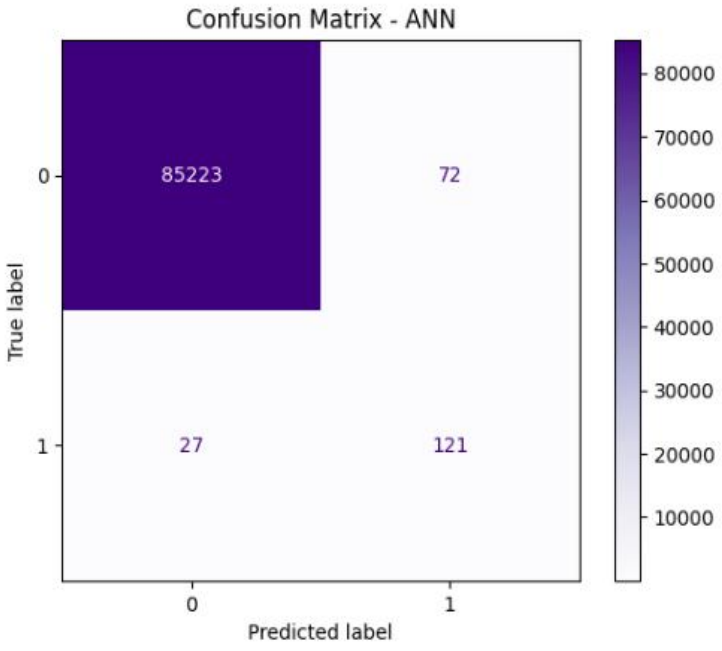

Figure 34Classification report and confusion matrix – ANN.

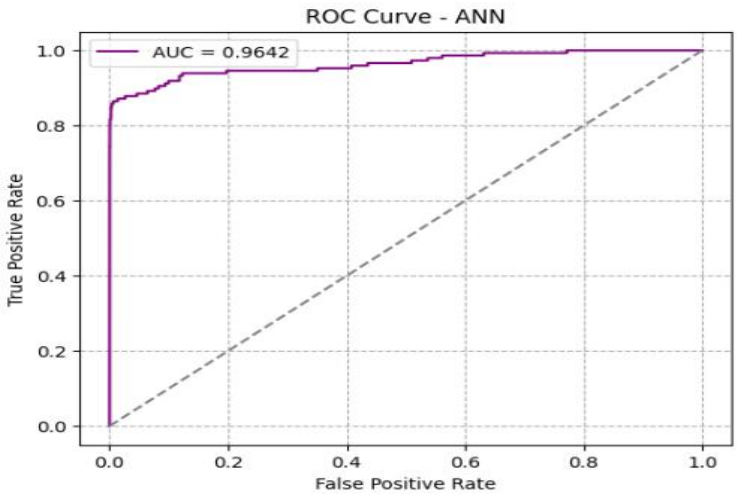

Figure 35: ROC curve – ANN.

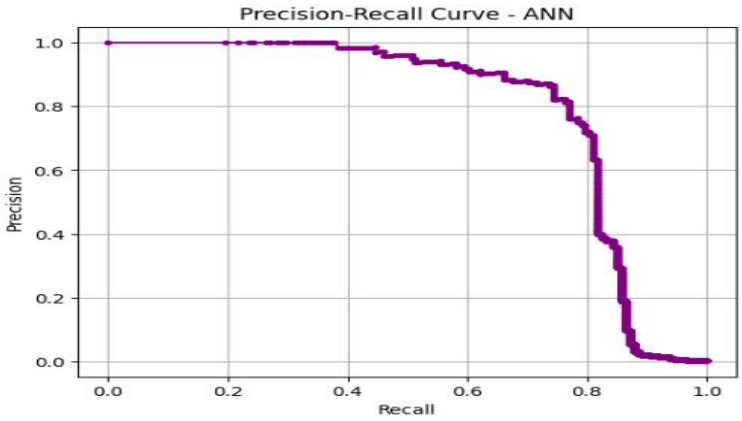

Figure 36: Precision-Recall curve – ANN.

Before hyperparameter tuning, all models were initially trained using their default configurations to establish baseline performance. The results provide a reference point to assess the impact of optimization techniques applied later. As shown in the comparison chart and summary table, Random Forest and XGBoost performed similarly well, achieving F1-scores of 0.8369 and 0.8194 respectively. The ANN yielded a moderate F1-score of 0.7097, while Logistic Regression significantly underperformed with an F1-score of just 0.2533. These differences highlight the importance of advanced ensemble and nonlinear models in handling complex fraud detection tasks. These observed trends are confirmed with evaluations based on the accuracy, AUC, precision, and recall metrics.

Table 8: Performance Comparison Across Models

Model	Accuracy	Precision (Fraud)	Recall (Fraud)	F1-Score (Fraud)	AUC
Logistic Regression	0.9983	0.1488	0.8514	0.2533	0.9665
Random Forest	0.9995	0.8806	0.7973	0.8369	0.9493
XGBoost	0.9994	0.8429	0.7973	0.8194	0.9779
ANN	0.9988	0.6269	0.8176	0.7097	0.9642

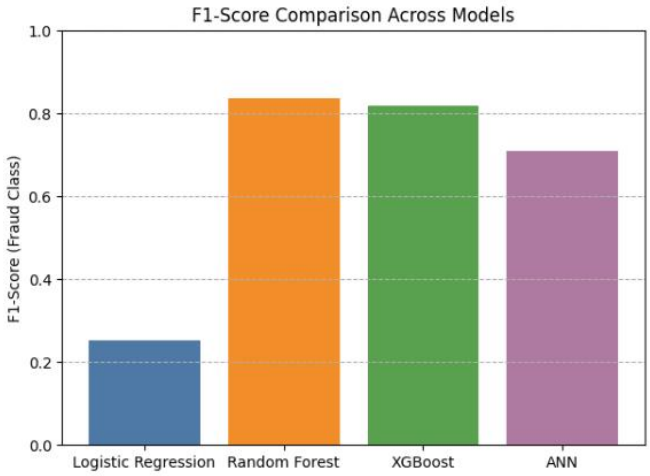

Figure 38: F1-Score Comparison before tuning Bar Plot.

4.5 Hyperparameter Tuning

Hyperparameter tuning is essential for optimizing ML model performance, especially when applied to sensitive and high-risk domains like fraud detection (Ali et al., 2023). Although these baseline models are good starting points, fine-tuning allows control parameters to be adjusted so that performance is maximised with measures such as precision, recall and F1-score. Considering the imbalanced nature of the dataset, and the objective of maximising efficiency of fraud detection, two models Random Forest and XGBoost were chosen as models for hyperparameter tuning, for the strength of their baselines and flexibility. For the tuning process RandomizedSearchCV was used which allows a time effective approach to parameter optimisation. Furthermore, cross-validation was also applied on Random Forest for robustness and stability of its results using varying subsets of the data.

4.5.1 RandomizedSearchCV Setup

In order to simplify the process of hyperparameter tuning, RandomizedSearchCV taken from the scikit-learn library was used. This method works especially well if a project involves a large parameter space because it samples a fixed number of parameters, rather than exhaustive evaluation of all combinations (Arafath et al., 2021). This method reduces considerably the computation time while conserving the potential to uncover optimal configurations. F1-score was the main tuning metric, which reflects the harmonic mean of precision and recall, both crucial for fraud detection, where cases of fraud detection (recall) and the prevention of false alarms (precision) are equally important.

The experiments were carried out using a portion of 50 000 samples taken from SMOTE-resampled training data to trade off feasibility of computing for learning capacity by the model. For Random Forest model parameters, the number of estimators, maximum tree depth, minimum samples required for splitting, and leaf node sizes were optimised. In the same way, for XGBoost parameters such as number of boosting rounds, learning rate, tree depth, subsampling ratio and column sampling ratio were investigated. These configurations were chosen because of already known impact on model complexity, overfitting control, learning stability.

4.5.2 Tuned Random Forest

All these after the tuning process, the best performing Random forest combination had 200 estimators, a maximum size of 20 for depth, minimum split size of 5, and a minimum of one sample per leaf. The model was re-trained with such parameters on the SMOTE-balanced training set and was tested on its original test set. The performance was significantly better than the baseline results. Specifically, the fraud class (Class 1) achieved a precision of 0.9009, recall of 0.6757, and F1-score of 0.7722, while the overall accuracy remained high at 99.93%.

The confusion matrix further illustrated the model's performance. Out of 148 fraudulent transactions, the model correctly identified 100 and misclassified 48. It also produced only 11 false positives from over 85,000 legitimate transactions. These outcomes demonstrate that the tuned Random Forest model became more conservative in its predictions better at avoiding false alarms, albeit slightly less aggressive in detecting every instance of fraud. This trade-off, common in imbalanced classification tasks, may be acceptable depending on the context, especially when the cost of false positives is high.

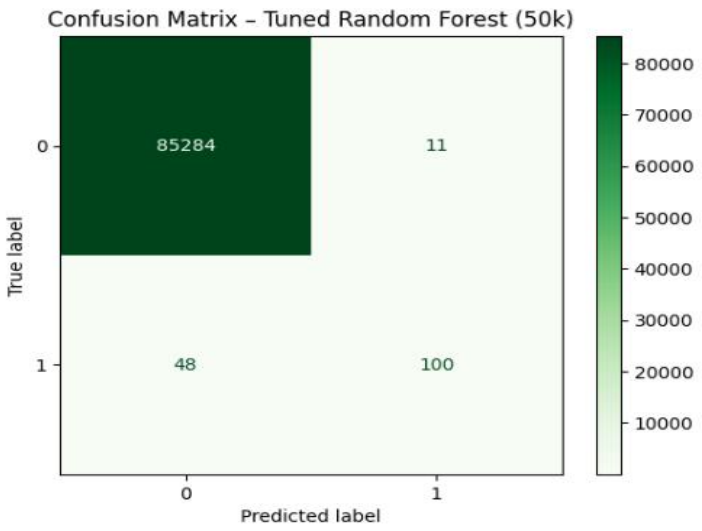

Figure 39: Random Forest evaluation after tuning.

4.5.3 Tuned XGBoost

XGBoost, known for its gradient boosting approach and high performance in structured datasets, also underwent tuning using RandomizedSearchCV. The best configuration included 200 estimators, a learning rate of 0.1, maximum depth of 5, and a row subsampling ratio of 0.8. Once retrained, the tuned XGBoost model demonstrated strong predictive performance. It achieved a precision of 0.9252, recall of 0.6689, and an F1-score of 0.7765 for the fraud class, with an overall accuracy of 99.94%.

Compared to the tuned Random Forest, XGBoost offered higher precision and a slightly higher F1-score, though at the cost of marginally lower recall. The confusion matrix reflected this shift, 99 fraudulent transactions were correctly identified, 49 were missed, and only 8 legitimate transactions were falsely labelled as fraud. These results suggest that the tuned XGBoost model is even more selective, effectively minimizing false positives—an important advantage in practical scenarios where every flagged transaction demands manual review or intervention.

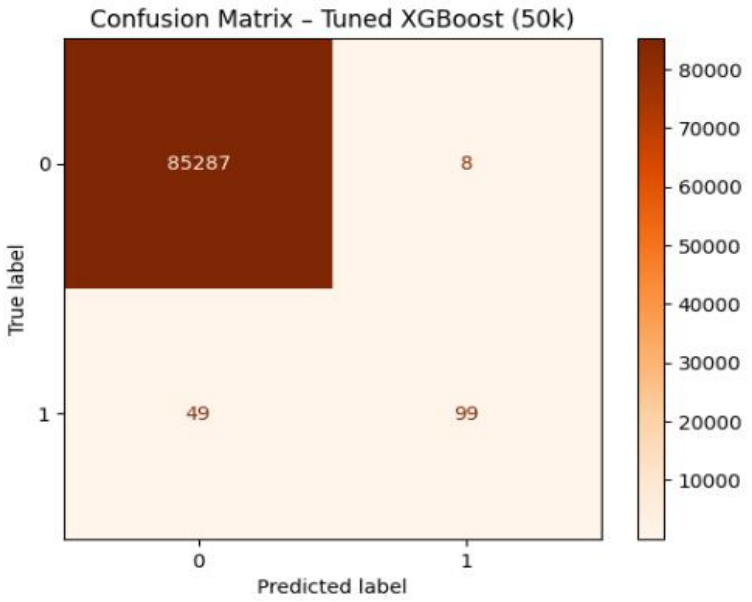

Figure 40: XGBoost evaluation after tuning.

4.5.4 Cross-Validation for Random Forest

To validate the consistency of the tuned Random Forest model, 5-fold cross-validation was conducted using the same 50,000-sample training subset. The F1-scores across the folds were as follows: 0.7586, 0.8462, 0.9333, 0.8387, and 0.7586, resulting in an average F1-score of 0.8271. These scores demonstrate that the model's performance is relatively stable across different data splits. While some variance occurred, likely due to the sparse presence of fraud cases in certain folds, the overall consistency reinforces confidence in the model's generalization capability. Notably, this cross-validation step was not repeated for XGBoost due to its higher computational cost. However, the consistent performance observed across evaluation metrics supports its reliability.

```
RF Cross-Validated F1 Scores: [0.75862069 0.84615385 0.93333333 0.83870968 0.75862069]
RF Mean CV F1 Score: 0.8270876472433759
```

Figure 41: Random Forest Cross-Validated F1 Scores.

The hyperparameter tuning process significantly enhanced the classification performance of both Random Forest and XGBoost models. For both, precision improved considerably, and overall F1-scores for the

fraud class increased post-tuning, even though a slight reduction in recall was observed. These shifts in performance indicate that the models became more confident and selective in flagging fraud, a common effect of tuning where false positives are minimized at the cost of a few missed detections. Between the two, XGBoost offered marginally better fraud classification metrics, making it ideal when false positives must be kept low. On the other hand, Random Forest showed stronger recall and more consistent performance across validation folds, making it a dependable choice in cases where missing fraudulent activity is more detrimental.

Figure 42: ML models' comparison after tuning.

4.6 Interpretation of Results

The comparative evaluation of the tuned ML models reveals that XGBoost performs slightly better than Random Forest in detecting credit card fraud, making it the most effective model in this study. After tuning, XGBoost achieved an F1-score of 0.7765 on the fraud class, marginally outperforming Random Forest, which scored 0.7722. Although both models delivered near-identical overall accuracy (0.9993), the key difference lies in how each model balances precision and recall on the minority (fraudulent) class, which is crucial in imbalanced classification problems. For the fraud class (label 1), XGBoost attained a precision of 0.9252 and recall of 0.6689, while Random Forest showed slightly higher precision (0.9009) but lower recall (0.6757). This means XGBoost had a marginally better ability to detect fraudulent cases without compromising too much on precision. In highly sensitive applications such as fraud detection,

recall often takes precedence because failing to detect fraud (false negatives) has more severe consequences than occasionally flagging a legitimate transaction (false positives). While both models maintained strong precision, XGBoost's slightly better recall and higher F1-score show that it more effectively balances both concerns.

It is also important to consider the macro average F1-score, which treats each class equally. XGBoost achieved a macro F1-score of 0.8881, compared to Random Forest's 0.8859. The difference may appear minor, but in operational terms, this could translate to many additional fraud cases being correctly detected over time, which is highly valuable in a real-world setting. Additionally, both models achieved nearly identical weighted averages (0.9993), reflecting their overall robustness, yet XGBoost does so with slightly greater efficiency on the minority class. Random Forest's higher precision is beneficial in scenarios where institutions want to reduce the number of false alarms and minimize manual review efforts. However, its lower recall could be problematic in systems where catching every possible fraud case is the top priority. On the other hand, XGBoost provides a better balance, it detects more fraud cases while still maintaining excellent precision, making it more suitable for environments where maximizing fraud coverage is critical.

Another advantage of XGBoost is its ability to regularize model complexity and prevent overfitting through built-in L1 and L2 penalties (Kabane, 2024). This is particularly valuable in fraud detection, where complex patterns often exist, and a model must generalize well to unseen transactions. While Random Forest is also a strong ensemble learner, its reliance on averaging many decision trees can sometimes limit its responsiveness to minority class patterns when not combined with strong tuning (Chogugudza, 2022). While both models show excellent performance, the tuned XGBoost model provides the best balance between detecting fraud (recall), minimizing false positives (precision), and maintaining stability across classes (F1-score). Therefore, it stands out as the most reliable and practically effective choice for credit card fraud detection in this research.

4.7 Importance of Feature Selection

Feature selection is a pivotal process in machine learning, particularly for tasks like fraud detection, where the goal is to identify rare, significant patterns within imbalanced datasets (Kajal & Kaur, 2021). In this study, feature importance was analysed for both Random Forest and XGBoost models, revealing the variables that were most influential in detecting fraudulent transactions. The results showed that V14 was by far the most important feature across both models, with V10 and V4 also emerging as significant in certain contexts. For Random Forest, the feature importance chart indicates that V14 has the highest value, far hitting other features. V10 came second in its importance to identify fraud transactions. The V17 was

was less ranked compared to V14 and V10. This implies that in Random Forest model, differentiation of fraud from those transactions that are not fraudulent depends heavily on V14 and V10, which are likely to capture special patterns of fraudulent behaviour. The relevance of these characteristics is further emphasised by their persistence in other models and their prominent role in fraud detection. The selection of most relating features, helps the model to concentrate on the data points that influence the results significantly, consequently increasing the model efficiency, minimising overfitting, and improving the interpretability of the model (Cheng, 2024).

XGBoost is even more V14-dominated, and V4 appears as the second-best feature. Ranking lower here is V10 while V17 has a moderate importance level. The difference in feature ranking between the two models implies that though V14 is the most important feature in both, XGBoost gives more weight on V4 while Random forest give more importance to V10. Such variance is probably the result of the different mechanisms that the two models use. Addressing errors is the focus of XGBoost's gradient boosting, causing extra importance in features such as V4 that might capture residuals from prior estimates more diligently. As a whole, the importance of V14 in the two models reveals its important contribution to fraudulent transaction identification. The variances in the rankings of other factors such as V10 V4 and V17 between the models indicate that feature selection needs to be individualised to the uniqueness of each algorithm. By concentrating their attention on these leading performing features both models are able to detect fraud better, at the same time, they are keeping their precision and false positives low. Therefore, feature selection not only enhances model performance but also enables the formulation of a system that allows the most influential factors to be used toward enhancing fraud detection.

4.8 Class Imbalance Impact

Class imbalance is a major problem of fraud detection, as fraudulent transactions form an extremely small minority part of the dataset (0.172% cases in the current study). This imbalance can turn out to be problematic because algorithms tend to make predictions based on majority class (legitimate transactions), which would often provide high false negatives (missed fraud cases). To address this, the Synthetic Minority SMOTE was used to balance the dataset by generating synthetic examples of the minority class (fraudulent transactions), thereby improving the model's ability to detect fraud. The application of SMOTE resulted in a balanced training dataset, as shown in the distribution before and after resampling. Before implementing the SMOTE, the data set contained 199820 legitimate transactions (Class 0) and only 344 fraudulent transactions (Class 1). After resampling the classes were appeared equally with 199,820 instances of legitimate transactions as well as 199,820 instances of fraud transactions. Such resampling enabled the models to learn decision boundaries between a legal and fraudulent transaction

better, which solved the class imbalance problem (Kerwin & Bastian, 2021). Improvement of model performance after SMOTE was clear. Random Forest and XGBoost both demonstrated an enhanced recall and general fraud detection, at the expense of neither their general accuracy. Although the models increased recall becoming better able to identify fraudulent cases, there was an increase in a small number of false positives. This balance between precision and recall is important in fraud detection systems, where there is increased tendency of shifting the scope of emphasis towards maximum fraud detection (high recall) at the expense of manual reviews caused by false positive.

Moreover, both models could discriminate between the two classes well after resampling based on the results of ROC and PR curves. The ROC-AUC scores were ideal for both models, with XGBoost having 0.9779 and Random Forest 0.9493 suggesting robust model performance in detecting fraud. PR curve further supported capability of the models to retain good precision and recall with excelling of XGBoost to balance both metrics. SMOTE made the detection of fraud possible to some reasonable extent for Random Forest and XGBoost from the imbalanced dataset. SMOTE enabled the models to exploit the synthetic samples generated from the minority class for better studies, improving recall and generally larger fraud detection performance. The trade-off between recall and precision, and the use of ROC and PR curves gave more attention to the models' ability to trade detection of fraud and avoiding false positives, a very important feature in practical fraud detection systems.

4.9 Model Reliability

Reliability of the model is critical in fraud detection, and wrong predictions may be costly both financially and reputationally. Random Forest and XGBoost were assessed to see whether they perform generalisation to new, unseen data to produce consistent performance over the long term. Although XGBoost and Random Forest produced good results, insight into their robustness versus fraud detection adds details that may guide the use of either as a model for actual implementation.

4.9.1 Cross-Validation Results and Performance Variability

To measure the reliability of Random Forest ethical 5-fold cross-validation was used. The F1-scores determined at the five folds varied between 0.7586 and 0.9333 with 0.8271 as the mean F1-score. This is as expected because of the greatly imbalanced nature of the data set. In some of the folds, the fewer number of fraudulent transactions made the task of detection harder, but Random Forest did show stable performance in total. This variability does mean that Random Forest can be sensitive to data distribution given that it matters when deploying those models in real-world settings where fraud cases can be different over time.

4.9.2 Stability and Generalization

Random Forest is a reliable an ensemble technique that develops various decision trees based on different subsets of data samples and averaging the results to cope with overfitting and enhance stability (Bakhtiari et al., 2023). This feature is useful in detecting fraud, since it is possible that patterns of fraud may change on a time basis. Although XGBoost has the higher complexity of the model and captures more complex, non-linear relationships by sequentially building trees it is more vulnerable to overfitting. However, by tuning its hyperparameters the risk is compensated, its generalisation ability is improved. These differences did not significantly affect the performance of both models as the latter continued to have high ROC-AUC scores. XGBoost performed with 0.9779 and Random Forest was 0.9493, a strong score to discriminate legitimate from fraudulent transactions even in imbalanced data.

4.9.3 Real-World Applicability

Adaptability to changing fraud patterns is important in fraud detection in that the model should adapt. Random Forest and XGBoost showed robust real-world generalisation capability. The XGBoost was slightly better than Random Forest in terms of precision and recall, in particular, detecting the subtle or developing patterns of fraud. The in-built regularisation techniques of XGBoost served to prevent overfitting making the model effective as the tactics of fraud changed. Conversely, Random Forest performed well in stability with low false positives and the avoidance of manual intervention, a critical aspect of fraud detection systems with low false alarm rates.

Both models were able to sustain reliable generalisation across differing fraud patterns, whereas XGBoost presented a more balanced precision, recall, while Random Forest delivered a more stable outcome. These results emphasise the fact that both models are reliable and robust they can be used for real-world fraud detection, and each one has complementary strengths depending on the needs in a specific application. Maximising of fraud detection is better fulfilled through XGBoost while Random Forest offers stability and precision in minimising false positives and manual reviews.

4.10 Real-World Application of Results

The choice of a proper fraud detection model is highly important as both Random Forest and XGBoost displayed significant performance, of which performance of each is advantageous. One of those depends on the requirements of the organisation while the other depends on the demands of the business. XGBoost performs better in terms of precision, which makes it the best for situations in which false positives have to be minimised at all cost. False positives that find valid transactions fraudulent produce avoidable

interventions, high investigation costs, and customer disappointment. In situations where customer experience and low false alarms are of concern, mere precision of XGBoost makes it the better option.

On the other hand, Random Forest excels in recall, which is essential for detecting as many fraudulent transactions as possible. Recall measures how effectively the model identifies actual fraud cases (Alfaiz & Fati, 2022). In high-stakes applications, missing fraud cases (false negatives) can result in significant financial losses and reputational damage. Therefore, Random Forest may be more suitable when it's vital to catch every fraudulent transaction, even at the cost of having a few more false positives. The decision between XGBoost and Random Forest ultimately hinges on the trade-off between false positives and false negatives, with XGBoost offering higher precision and Random Forest delivering stronger recall. In real-world fraud detection, practical considerations like the cost of false positives and false negatives play a crucial role. False positives incur additional costs due to manual reviews, while false negatives allow fraud to go undetected, leading to potential financial losses (Vorobyev & Krivitskaya, 2022). Organizations need to balance these costs carefully. If minimizing undetected fraud is a priority, Random Forest is more fitting due to its better recall. However, in environments focused on reducing manual review efforts and minimizing false alarms, XGBoost's higher precision offers significant advantages.

Both models also require substantial integration with current fraud detection systems to process transactions in real or near real times. The operational cost of XGBoost is higher because of gradient boosting mechanism, but it is highly efficient provided that it is tuned effectively (Bentéjac et al., 2021). Random Forest which is more effective might be more appropriate in resource-limited settings or when real-time detection is not so important. Besides, regular training and regular model updating is required because with the passage of time, the fraud patterns change. Interpretability of the model is also an important consideration, especially where the model is used in regulated industries. At most times Random Forest provides higher interpretability which comes in handy if regulatory explanations of decisions are needed (Dube & Verster, 2024). Although XGBoost calculates the metrics for the feature importance, its mechanism of decision making is more complicated, which may constrain the element of transparency in some situations. Thus, Random Forest could be better suited to the settings where the explanations for automated decisions are critical.

4.11 Limitations & Future Enhancements

Although, Random forest and XGBoost methods used in current study represented strong results in recognising credit card frauds several limitations must be addressed to enhance their effectiveness and applicability in credit card fraud detection models used in real world. The limitations of the current data, processes used to tune the models, and computational resources are the sources of these limitations, and

these will provide for the basis of improvements on fraud detection models in the future. One major limitation of this study was the tuning constraints especially in terms of time and hardware. Although both Random Forest and XGBoost have been tuned using methods such as RandomizedSearchCV, the size of the search space for possible hyperparameters to test was unavoidably constrained because of the time cost of running exhaustive searches of all possible combinations (Tiwari et al., 2022). This restricted the possibility of proper optimization of the models and possibly led to sub-optimal hyperparameter settings for models. In practical use, where the price for false positives and false negatives is high, deeper tuning using more computational resources may result in an additional level of precision for the models. In future work, Bayesian optimization, or a GridSearchCV with a sharper search space could be used to find better performing configurations of hyperparameters.

The second limitation was the absence of cross-validations of XGBoost. Although random forest received a very strict validation using 5-fold cross-validation, only the XGBoost regressor went through a basic train-test split. Cross-validation is very important in the estimation of the stability of a model and to the conservation of consistency as a model works on different subsets of data. Without such an extra validation, the performance metrics of XGBoost may be unable to indicate its full potential in managing unseen data. In further studies, there is need to perform the same cross validation testing in order to validate the degree of generalizability, and to minimise overfitting. In addition, the models in this study were based on an anonymized dataset of credit card transactions and as effective as a tool for testing, might not reflect the nuances of the real fraud patters. Through this dataset, more contextual information is lacking such as customer demographics, merchant details, or time-based transactions history which would enrich the probability of fraud detection by affording a better view of the transactions.

Featuring engineering with the help of domain specific knowledge may help to include these missing variables which might enhance the models' capacity to detect fraud. One of the ways in which future improvements can be performed is experimenting with the techniques of synthetic data generation or using external sources of data that can be used to supplement it with the necessary context. The future work could analyse how explainability tools such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) can increase the models' interpretability level (Ahmad et al., 2024). Such tools may show how specific predictions make sense, which is especially important for financial institutions where transparency is necessary for a regulatory compliance. Application of such tools would give clearer insights on how the models are making decisions – easier to justify flagged transactions and win a customer's confidence in the system.

4.12 Chapter conclusion

When both XGBoost and Random Forest models are compared, then XGBoost emerges as the better model to use in the process of credit card fraud detection. Although both models had near identical overall accuracy (0.9993), XGBoost outperformed the other one when assessed against F1-score and recall. XGBoost achieved 66.89% recall relative to 67.57% of Random Forest which means that XGBoost was efficient in detecting fraudulent transactions at a good precision 92.52%. In fraud detection, recall is usually given much more import than precision, since it entails less grievous consequences to miss a fraud (false negatives), than mistaking a true transaction into a fraud (false positives).

A further confirmation of the efficiency of XGBoost is given by the macro average F1-score, which is 0.8881, only slightly higher than the 0.8859 from Random Forest. Although difference might seem marginal, it is translated into feasibility of detection of more fraudulent cases on practical performance, where such activity is imperative for losses minimization. And both models scored almost equal weighted averages, which reflects their robustness overall, only XGBoost outperformed demonstrating better efficiency when handling the minority class. Another advantage of XGBoost is that it has incorporation of built-in regularisation (L1 and L2 penalties), which helps for underlying the issue of overfitting and drives up generalisation of data not seen before. This is especially useful in the detection of fraud where patterns may be so complex and ever changing. Although, Random Forest had slightly greater precision, its recall was lower which means it is likely to overlook vital fraud cases. XGBoost shows the better balance between recall, precision, generalisation and the best option which is reliable and effective in credit card fraud detection.

Chapter 5 Conclusion

5.1 Chapter Introduction

This chapter reports the conclusions made based on the research of detecting credit card fraud via ML methods. It seeks to synthesise the main findings of the study, provide practical implications of these results for the financial industry, discuss limitations encountered during the research process, and recommend guidelines on how to move forward in subsequent research.

The study investigated the application of ML methods, such as Logistic Regression, Random Forest, XGBoost and Neural Networks to identify fraudulent credit card transactions. The key object of the evaluation was the performance of these models as Fraud detection models on imbalanced datasets, with the particular attention paid to such metrics as precision, recall, F1-score. The chapter will review the means these models may be used in real-life fraud detection systems, difficulties with model deployment, and recommendations to enhance fraud detection techniques with the help of advanced ML techniques and better data.

5.2 Summary of Key Findings

This research was conducted to assess the performance of ML models for recognising credit card fraud with a particular emphasis on a comparison of Logistic Regression, Random Forest, XGBoost, and ANN. The main findings showcase the model's work on imbalanced datasets, the need for proper evaluation metrics, and effect of imbalanced data on fraud detection.

5.2.1 Effectiveness of ML Models

XGBoost resulted to be the best model to be used in a fraud detection system; higher in precision, recall and F1-score than the other algorithms. This group of methods uses gradient boosting allowing it to learn from the mistakes of previous models strengthening its potential to detect fraudulent transactions, which is especially significant because of their rarity. XGBoost managed to produce a better performance due to its capability to handle complicated patterns of data. Random Forest did well, but it became a little less efficient compared to XGBoost in detecting frauds. As an aggregation of trees, it combines results of many trees and increases predictability. Although Random Forest is resilient to overfitting and manages nonlinear data properly its performance with regard to fraud detection precision was inferior to XGBoost.

Logistic Regression while cheap and interpretable failed in the complexity of fraud detection on imbalanced datasets. It did well on identifying linear relationships but was unable to capture the complex nonlinear patterns inherent of fraudulent activity, so, it became a less desirable option for this endeavour.

ANN had potential, but were computationally expensive and training needed for it was long. Even though deep learning models are capable of modelling complex interactions, but opaqueness and time-consuming training prevented practical real time application for fraud detection. In addition, the requirement for interpretability in the financial systems reduced the appeal of ANN in the case of fraud detection in this context.

5.2.2 Impact of Data Imbalance

Class distribution imbalance is one of the key problems of fraud detection (Kalid et al., 2024). The fraudulent transactions comprising only a very small fraction of the total data set. Such an imbalance makes it hard for traditional models to efficiently verify fraud because they are prone to prediction bias of majority class (non-fraudulent transactions).

In order to overcome this, the study employed methods, such as SMOTE and under-sampling to balance the dataset. SMOTE created synthetic samples of fraudulent transactions, and the models gained from these scarce instances. Such methods proved successful in enhancing model performance, with particular success on fraud detection, and model bias towards non-fraudulent transactions.

5.2.3 Importance of Evaluation Metrics

The similarity of performance estimates across fraud modelling settings led to the recommendation to use precision, recall, and F1-score in the model's performance assessment, rather than accuracy. In an imbalanced dataset, accuracy may turn out to be misleading-in such cases a model would be able to declare all transactions legitimate and still obtain a high amount of accuracy because of the high numbers of non-fraudulent transactions.

Precision and recall were important in this study because precision denotes the portion of fraud cases which are properly identified, and recall measures the ability of the model to detect all fraudulent transactions. F1-score, as the metric of precision and recall balance, was especially helpful with regard to assessing the models' overall performance. On these metrics, XGBoost and Random Forest did well, confirming their applicability in fraud detection in finance-related datasets.

5.3 Practical Implications

The results of the work have great practical value for the financial industry, especially in the design and implementation of fraud detection systems. As more digital transactions are recorded and as the cases of credit card fraud continue to increase, financial institutions need also to implement more sophisticated

technologies, with a view to preventing fraudulent activities (Udeh et al., 2024). The research shows how ML models (XGBoost and Random Forest) can be used to enhance fraud detection functionality in real-world settings. These models offer an increased number of the identified fraudulent transactions compared to the rule-based systems that are static and can hardly change with time due to changing fraud patterns. Results of the research suggest that ML systems can provide a more adaptive, real-time and accurate fraud duplicate detection using historical data and the ability to adapt their decision-making process.

The major benefit of ML models, including XGBoost and Random Forest, is their capability to process large and complicated datasets involves many features, most of which remain undiscoverable for the human or conventional approaches (Fatima et al., 2023). These models are also efficient in uncovering gentle patterns and anomalies in transactional data, which is important to reveal advanced schemes of fraud that can escape from traditional detection systems. For financial institutions, this means a system that not only improves the accuracy of fraud identification but also decreases the chance of false positives, whereby real transactions are often falsely assumed to be fraudulent. This increased precision results in improved customer satisfaction because the customers hardly experience interruptions of their financial operations like transaction rejections or freeze of accounts which is mostly triggered by the rule-based systems' false alarms.

However, the application of ML models in operations is not straightforward. The outcome of this study also calls for strong computational infrastructure in support of these models especially to be used in real-time fraud detection systems. Training and tuning these models especially deep learning models will require a lot of computations. Financial institutions need also to ensure that their systems have the capacity to process massive volumes of transactional data in real time without sacrificing on performance (Adeleke et al., 2022). Although ML may be able to help improve the detection of newly developing and evolving fraud techniques, it should operate collaboratively with conventional approaches in order to develop a more comprehensive fraud prevention plan. This hybrid method enables joining the advantages of both systems. ML's capability of detecting delicate patterns and a capacity to adjust new fraud behaviours and rule-based systems' capacity of observing for known fraud features.

5.4 Limitations of the Study

Though promising results were achieved in the course of this study, a number of limitations should be recognized. Such shortcomings may affect the generalizability of findings, and stimulate areas for further investigation to improve fraud detection systems. The problems with the dependent ML models represent one of the main limitations of this study. Although models like XGBoost and Random Forest did well,

they are not constraint free. For instance, XGBoost even with its superior performance, is computationally latent and may not be able to perform optimally in environments that experience a resource constraint (Noviandy et al., 2023). In the same way, although Random Forest is strong and reliable in dealing with imbalanced datasets, it is not free from the problem of overfitting, especially, when the number of trees in generation is excessive. This can lead to models which are too specific to the training data and less able to generalize to new, unseen data. Further, XGBoost and Random Forest do not possess the readability that is needed in many financial institutions. Transparency is important in financial sectors where one needs to understand why a decision on fraud detection is made for regulatory purposes and customer satisfaction (Ejiofor, 2023). Models such as, ANN that can provide even more favorable performance, are many times specified as the black box type, less explainable and deployable for real-world financial systems.

The limitation of the second type is due to what dataset was used in this study. The dataset used was an anonymised Kaggle dataset which despite its widespread use for fraud detection research, comes with an inbuilt set of constraints. For example, it is not rich in contextual information which may give deeper insights of the behaviour of the cardholders. Some of the potential features that may increase the model's performance are customer demographics, the history of spending or behavioural signals. In addition, real-world transaction data is constantly being updated and fraud tactics change. These dynamic factors were not considered in the study, i.e. the models are being trained on static data that may not capture the changes in fraud schemes. Finally, techniques such as SMOTE were used to balance the dataset, however, such approaches may not produce most realistic fraudulent data at all times. Synthetic data could not reflect the full spectrum of fraud type, which may result in a model that is less flexible to new fraud strategies in the operational settings. Although the research exhibits promising potential in the use of ML for fraud detection, additional research and future developments in both techniques of models and quality of data are required to overcome these limitations.

5.5 Ethical Considerations and Data Privacy

Ethical issue and data privacy play a vital role when using ML models for finding credit card frauds. Financial institutions work with enormous amounts of sensitive personal and financial information that should be secured following such regulations as the GDPR and PCI DSS. A major ethical issue is to guarantee customer data anonymization and security. In the current study, anonymised datasets were used but in the real use cases the management of personal data should follow strict privacy regulations to prevent identity thefts or unauthorised access to data. Financial institutions should deploy private information-preserving measures like data encryption, tokenization to protect delicate data.

Algorithmic fairness and transparency are yet another ethical issue. ML models, especially those, that are applied to fraud detection, may seem to be “black boxes” because of their complexity and lack of metaphoric transparency (Pedreschi et al., 2019). For financial services, it is also important that both customers and regulators can understand why a specific transaction was marked as fraudulent. Achieving transparency with regard to model decisions and create explainable-AI is crucial to uphold trust and compliance obligations of regulation. Similarly, attempts should be made to prevent the bias of the models that may discriminate certain groups of people along their demographic lines such as location or spending habit in order for all customers to be treated fairly.

5.6 Future Research Directions

Although this study has identified insights on ML use for credit card fraud detection, there are a number of areas, which require more investigation if the model is to be made robust and presumed challenges addressed. Integration of advanced ML models, specifically deep learning practices, is one of the main directions. The complex patterns in the data could have been demonstrated by neural networks such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for better possibilities (Banu et al., 2024). These models are able to learn more complex fraud behaviour and may enhance detection rates especially for complex fraud scheme. However, the trade-off is greater computational costs, as well as the high requirement of data, which are also challenges that need to be addressed.

Another area that would show additional potential for future research is the integration of unstructured and behavioural data. This study was based on transactional data that are structured, but real-world fraud detection systems will see great value in incorporating information from various diverse sources, such as social media, geolocation, and user behaviour data. This non-structured data might offer more context, and also detect fraudulent patterns that would have been overlooked. Additionally, incorporating behavioural analytics that track how a user interacts with their accounts such as, login patterns, IP addresses, device usage, could improve the robustness of fraud detection systems (Chakraborty et al., 2022).

Real-time adaptive fraud detection is another critical avenue for future research. Fraud tactics evolve constantly, so systems need to adapt quickly. Implementing models that use reinforcement learning or online learning could help detect new fraud schemes in real-time and update fraud detection models autonomously as new transaction data comes in. Lastly, model explainability and fairness remain key concerns. Future research should focus on improving model transparency to ensure compliance with regulations and build trust with consumers. Developing explainable AI will help mitigate the “black-box”

nature of certain models, enabling financial institutions to offer more transparent fraud detection processes.

This study has demonstrated the effectiveness of ML techniques, specifically XGBoost and Random Forest, in detecting credit card fraud. By comparing various models, it was evident that ensemble methods, such as XGBoost, offer superior performance in identifying fraudulent transactions, particularly in imbalanced datasets. The research highlighted the importance of using appropriate evaluation metrics, such as precision, recall, and F1-score, instead of relying solely on accuracy, which can be misleading in fraud detection scenarios. While ML models significantly outperformed traditional rule-based systems, challenges related to computational complexity, data imbalance, and model interpretability remain. The study also emphasized the need for real-time adaptation of fraud detection systems and the potential of incorporating unstructured and behavioural data to further improve detection capabilities. Moving forward, research should focus on advanced deep learning techniques, model transparency, and adaptive systems to address the evolving nature of financial fraud and improve overall security.

References

Aaron, W. C., Irekponor, O., Aleke, N. T., Yeboah, L., & Joseph, J. E. (2024). Ma-chine learning techniques for enhancing security in financial technology systems.

Abbasi, M., & Shah, M. A. (2022, June). Credit card fraud detecting using MLclassifiers in stacking ensemble technique. In IET Conference Proceedings CP801 (Vol. 2022, No. 8, pp. 76-81). Stevenage, UK: The Institution of Engineering and Technology.

Matsuk, S. (2022, December 6). Fraud Detection in E-Commerce. Retrieved June 23, 2025, from Amazinum website: https://amazinum.com/insights/fraud-detection-in-e-commerce/

Adeleke, A. G., Sanyaolu, T. O., Efunniyi, C. P., Akwawa, L. A., & Azubuko, C. F. (2022). Optimizing systems integration for enhanced transaction volumes in Fintech. Finance & Accounting Research Journal P-ISSN, 345-363.

Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., ... & Eshun, J. (2023). A supervised MLalgorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6, 100163.

Agarwal, A., & Ratha, N. K. (2021, November). Black-Box Adversarial Entry in Finance through Credit Card Fraud Detection. In CIKM Workshops.

Aghware, F. O., Ojugo, A. A., Adigwe, W., Odiakaose, C. C., Ojei, E. O., Ashioba, N. C., ... & Geteloma, V. O. (2024). Enhancing the random forest model via synthetic minority oversampling technique for credit-card fraud detection. Journal of Computing Theories and Applications, 1(4), 407-420.

Ahammad, J., Hossain, N., & Alam, M. S. (2020, January). Credit card fraud detection using data pre-processing on imbalanced data-Both oversampling and undersampling. In Proceedings of the International Conference on Computing Advancements (pp. 1-4).

Ahmad, T., Katari, P., Pamidi Venkata, A. K., Ravi, C., & Shaik, M. (2024). Explainable AI: Interpreting Deep Learning Models for Decision Support. Advances in Deep Learning Techniques, 4(1), 80-108.

Ahmadi, S. (2023). Open AI and its impact on fraud detection in financial industry. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (online), 2(3), 263-281.

Ahmed, J., & Green II, R. C. (2022). Predicting severely imbalanced data disk drive failures with MLmodels. MLwith Applications, 9, 100361.

Ahmed, M., Ansar, K., Muckley, C. B., Khan, A., Anjum, A., & Talha, M. (2021). A semantic rule based digital fraud detection. PeerJ Computer Science, 7, e649.

Alamri, M. and Ykhlef, M. (2022a). Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques. Electronics, [online] 11(23), p.4003. doi:https://doi.org/10.3390/electronics11234003.

Alenzi, H. Z., & Aljehane, N. O. (2020). Fraud detection in credit cards using logistic regression. International Journal of Advanced Computer Science and Applications, 11(12).

Alfaiz, N. S., & Fati, S. M. (2022). Enhanced credit card fraud detection model using machine learning. Electronics, 11(4), 662.

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., ... & Saif, A. (2022). Financial fraud detection based on machine learning: a systematic literature review. Applied Sciences, 12(19), 9637.

Ali, Y. A., Awwad, E. M., Al-Razgan, M., & Maarouf, A. (2023). Hyperparameter search for MLalgorithms for optimizing the computational complexity. Processes, 11(2), 349.

Alkattab, Y., & Edén Wallberg, O. (2024). Comparative Analysis of MLand Deep Learning Models for Card Fraud Detection.

Alzoubi, H. M., & Ghazal, T. M. (2022). The effect of e-payment and online shopping on sales growth: Evidence from banking industry. International Journal of Data and Network Science, 6(4), 1369-1380.

Amin, F., & Mahmoud, M. (2022). Confusion matrix in binary classification problems: A step-by-step tutorial. Journal of Engineering Research, 6(5), 0-0.

Arafath, Y., Roy, A. C., Shamim Kaiser, M., & Arefin, M. S. (2021, December). Developing a framework for credit card fraud detection. In Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021 (pp. 637-651). Singapore: Springer Singapore.

Arambawela, M., & Aponso, A. (2024, February). Using MLto Identify and Categorize Personally Identifiable Information and Payment Card Industry Data in Textual Content. In 2024 4th International Conference on Advanced Research in Computing (ICARC) (pp. 201-205). IEEE.

Arshad, M., Farhaoui, Y., & Shamim, R. (2024, April). Optimizing Hyperparameters for Fraud Detection: A Comparative Analysis of MLAlgorithms. In The International Workshop on Big Data and Business Intelligence (pp. 218-228). Cham: Springer Nature Switzerland.

Babu, A. M., & Pratap, A. (2020, December). Credit card fraud detection using deep learning. In 2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS) (pp. 32-36). IEEE.

Bakhtiari, S., Nasiri, Z., & Vahidi, J. (2023). Credit card fraud detection using ensemble data mining methods. Multimedia Tools and Applications, 82(19), 29057-29075.

Banu, S. R., Gongada, T. N., Santosh, K., Chowdhary, H., Sabareesh, R., & Muthuperumal, S. (2024, April). Financial fraud detection using hybrid convolutional and recurrent neural networks: An analysis of unstructured data in banking. In 2024 10th International Conference on Communication and Signal Processing (ICCSP) (pp. 1027-1031). IEEE.

Beju, D. G., & Făt, C. M. (2023). Frauds in banking system: Frauds with cards and their associated services. In Economic and financial crime, sustainability and good governance (pp. 31-52). Cham: Springer International Publishing.

Bello, H. O., Ige, A. B., & Ameyaw, M. N. (2024). Adaptive MLmodels: concepts for real-time financial fraud prevention in dynamic environments. World Journal of Advanced Engineering Technology and Sciences, 12(02), 021-034.

Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & Egwuonwu, M. N. (2023). Analysing the impact of advanced analytics on fraud detection: a MLperspective. European Journal of Computer Science and Information Technology, 11(6), 103-126.

Bello, O. A., Ogundipe, A., Mohammed, D., Adebola, F., & Alonge, O. A. (2023). AI-Driven Approaches for real-time fraud detection in US financial transactions: challenges and opportunities. European Journal of Computer Science and Information Technology, 11(6), 84-102.

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54, 1937-1967.

Bharadwaj, C. (2021). Financial Fraud Detection Using Machine Learning: A Comprehensive Guide. [online] Appinventiv. Available at: https://appinventiv.com/blog/role-of-machine-learning-in-financial-fraud-detection/ [Accessed 20 Jun. 2025].

Borwell, J., Jansen, J., & Stol, W. (2022). The psychological and financial impact of cybercrime victimization: A novel application of the shattered assumptions theory. Social Science Computer Review, 40(4), 933-954.

Brandt, J., & Lanzén, E. (2021). A comparative review of SMOTE and ADASYN in imbalanced data classification.

Btoush, E. A. L. M., Zhou, X., Gururajan, R., Chan, K. C., Genrich, R., & Sankaran, P. (2023). A systematic review of literature on credit card cyber fraud detection using machine and deep learning. PeerJ Computer Science, 9, e1278.

Carcillo, F., Le Borgne, Y. A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. Information sciences, 557, 317-331.

Chakraborty, D., Paul, A., & Kaur, G. (2022). Microeconomics: machine learning model with behavioural intelligence to reduce credit card fraud. International Journal of Electronic Banking, 3(4), 358-378.

Chaquet-Ulldemolins, J., Gimeno-Blanes, F. J., Moral-Rubio, S., Muñoz-Romero, S., & Rojo-Álvarez, J. L. (2022). On the black-box challenge for fraud detection using ML(I): Linear models and informative feature selection. Applied Sciences, 12(7), 3328.

Cheng, X. (2024). A Comprehensive Study of Feature Selection Techniques in MLModels. Insights in Computer, Signals and Systems, 1(1), 10-70088.

Chogugudza, M. (2022). The classification performance of ensemble decision tree classifiers: A case study of detecting fraud in credit card transactions. Identifier: vital, 69317.

Cruz, M. V., Usha, G., Vinoth, N. A. S., Anbarasi, A., Thamizhamuthu, R., & Gautam, K. (2024, October). A Comparative Technique with Credit card based Fraud detection with MLand Quantum Algorithm. In 2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS) (pp. 408-414). IEEE.

Dickerson, P., & Worthen, J. (2024, May). Optimizing Pipeline Systems for Greater Precision, Efficiency & Safety Using Emerging Technologies. In PSIG Annual Meeting (pp. PSIG-2426). PSIG.

Dileep, M. R., Navaneeth, A. V., & Abhishek, M. (2021, February). A novel approach for credit card fraud detection using decision tree and random forest algorithms. In 2021 Third International

Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1025-1028). IEEE.

Dube, L., & Verster, T. (2024). Interpretability of the random forest model under class imbalance. Data Science in Finance and Economics, 4(3), 446-468.

Ejiofor, O. E. (2023). A comprehensive framework for strengthening USA financial cybersecurity: integrating machine learning and AI in fraud detection systems. European Journal of Computer Science and Information Technology, 11(6), 62-83.

Fanai, H., & Abbasimehr, H. (2023). A novel combined approach based on deep Autoencoder and deep classifiers for credit card fraud detection. Expert Systems with Applications, 217, 119562.

Fatima, S., Hussain, A., Amir, S. B., Ahmed, S. H., & Aslam, S. M. H. (2023). XGBoost and random forest algorithms: an in depth analysis. Pakistan Journal of Scientific Research, 3(1), 26-31.

Genc, B., & Tunc, H. Ü. S. E. Y. İ. N. (2019). Optimal training and test sets design for machine learning. Turkish Journal of Electrical Engineering and Computer Sciences, 27(2), 1534-1545.

Geng, Y., Li, Q., Yang, G., & Qiu, W. (2024). Logistic regression. In Practical MLIllustrated with KNIME (pp. 99-132). Singapore: Springer Nature Singapore.

Georgieva, S., Markova, M., & Pavlov, V. (2019, October). Using neural network for credit card fraud detection. In AIP Conference Proceedings (Vol. 2159, No. 1). AIP Publishing.

Ghaleb, F. A., Saeed, F., Al-Sarem, M., Qasem, S. N., & Al-Hadhrami, T. (2023). Ensemble synthesized minority oversampling-based generative adversarial networks and random forest algorithm for credit card fraud detection. IEEE Access, 11, 89694-89710.

Gowda, C. (2022). Understanding Fraud Risk in E-Commerce with Special Emphasis on Credit Card Fraud and Triangulation Fraud. Issue 6 Indian JL & Legal Rsch., 4, 1.

Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2022). Fraud detection in banking data by MLtechniques. Ieee Access, 11, 3034-3043.

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: a review of anomaly detection techniques and recent advances. Expert systems With applications, 193, 116429.

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud

Jain, A., Purwar, A., & Yadav, D. (2021). Credit card fraud detection using k-means and fuzzy c-means. In Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies (pp. 216-240). IGI Global.

Jain, R. (2023). K-Means Clustering: Use Cases, Advantages and Working Principle. [online] Bombaysoftwares.com. Available at: https://www.bombaysoftwares.com/blog/introduction-to-k-means-clustering [Accessed 20 Jun. 2025].

Juusola, K., Boakye, K. G., Blankson, C., & Cao, G. (2023). A comparative examination of the motivating factors underpinning consumers' loyalty toward credit card usage in the United States and France. International Journal of Bank Marketing, 41(7), 1743-1768.

Jůzová, A. (2024). Comparative aAnalysis of Unsupervised Anomaly Detection Methods for Credit Card Fraud Detection.

Kabane, S. (2024). Impact of Sampling Techniques and Data Leakage on XGBoost Performance in Credit Card Fraud Detection. arXiv preprint arXiv:2412.07437.

Kajal, D., & Kaur, K. (2021). Credit card fraud detection using imbalance resampling method with feature selection. Int. J, 10(3).

Kalid, S. N., Khor, K. C., Ng, K. H., & Tong, G. K. (2024). Detecting frauds and payment defaults on credit card data inherited with imbalanced class distribution and overlapping class problems: A systematic review. IEEE Access, 12, 23636-23652.

Kamalaruban, P., Pi, Y., Burrell, S., Drage, E., Skalski, P., Wong, J., & Sutton, D. (2024, November). Evaluating Fairness in Transaction Fraud Models: Fairness Metrics, Bias Audits, and Challenges. In Proceedings of the 5th ACM International Conference on AI in Finance (pp. 555-563).

Kaur, H., Pannu, H. S., & Malhi, A. K. (2019). A systematic review on imbalanced data challenges in machine learning: Applications and solutions. ACM computing surveys (CSUR), 52(4), 1-36.

Kerwin, K. R., & Bastian, N. D. (2021). Stacked generalizations in imbalanced fraud data sets using resampling methods. The Journal of Defense Modeling and Simulation, 18(3), 175-192.

Khaled, A. A., Hasan, M. M., Islam, S., Papastergiou, S., & Mouratidis, H. (2024, June). Synthetic Data Generation and Impact Analysis of MLModels for Enhanced Credit Card Fraud Detection. In IFIP International Conference on Artificial Intelligence Applications and Innovations (pp. 362-374). Cham: Springer Nature Switzerland.

Khalid, A. R., Owoh, N., Uthmani, O., Ashawa, M., Osamor, J., & Adejoh, J. (2024). Enhancing credit card fraud detection: an ensemble MLapproach. Big Data and Cognitive Computing, 8(1), 6.

Khalid, A.R., Nsikak Owoh, Omair Uthmani, Ashawa, M., Osamor, J. and Adejoh, J. (2024). Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach. Big Data and Cognitive Computing, [online] 8(1), pp.6–6. doi:https://doi.org/10.3390/bdcc8010006.

Khanday, S. A., Vali, K. M., Junaid, M., & Ahmad, M. (2024). Understanding Positivism: A Qualitative Exploration of Its Principles and Relevance Today. European Journal of Applied Sciences–Vol, 12(6).

Khare, P., & Srivastava, S. (2023). AI-Powered Fraud Prevention: A Comprehensive Analysis of MLApplications in Online Transactions. J. Emerg. Technol. Innov. Res, 10(9), 2349-5162.

Khatri, S., Arora, A., & Agrawal, A. P. (2020, January). Supervised MLalgorithms for credit card fraud detection: a comparison. In 2020 10th international conference on cloud computing, data science & engineering (confluence) (pp. 680-683). IEEE.

Khurana, R. (2020). Fraud detection in ecommerce payment systems: The role of predictive ai in real-time transaction security and risk management. International Journal of Applied MLand Computational Intelligence, 10(6), 1-32.

Kilickaya, O. (2024). Credit Card Fraud Detection: Comparison of Different MLTechniques. International Journal of Latest Engineering and Management Research (IJLEMR), 9(2), 15-27.

Kolevski, D., Michael, K., Abbas, R., & Freeman, M. (2022, November). Cloud computing data breaches in news media: Disclosure of personal and sensitive data. In 2022 IEEE International Symposium on Technology and Society (ISTAS) (Vol. 1, pp. 1-11). IEEE.

Kumar, J., & Saxena, V. (2022). Rule-based credit card fraud detection using user's keystroke behavior. In Soft Computing: Theories and Applications: Proceedings of SoCTA 2021 (pp. 469-480). Singapore: Springer Nature Singapore.

Kumar, P., & Iqbal, F. (2019, April). Credit card fraud identification using MLapproaches. In 2019 1st International conference on innovations in information and communication technology (ICIICT) (pp. 1-4). IEEE.

Kumar, S., Gunjan, V. K., Ansari, M. D., & Pathak, R. (2022). Credit card fraud detection using support vector machine. In Proceedings of the 2nd International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications: ICMISC 2021 (pp. 27-37). Springer Singapore.

Leevy, J. L., Hancock, J., & Khoshgoftaar, T. M. (2023). Comparative analysis of binary and one-class classification techniques for credit card fraud data. Journal of Big Data, 10(1), 118.

Madhurya, M. J., Gururaj, H. L., Soundarya, B. C., Vidyashree, K. P., & Rajendra, A. B. (2022). Exploratory analysis of credit card fraud detection using MLtechniques. Global Transitions Proceedings, 3(1), 31-37.

Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M. S., & Zeineddine, H. (2019). An experimental study with imbalanced classification approaches for credit card fraud detection. Ieee Access, 7, 93010-93022.

Maram Alamri and Mourad Ykhlef (2022b). Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques. Electronics, [online] 11(23), pp.4003–4003. doi:https://doi.org/10.3390/electronics11234003.

Mienye, I. D., & Jere, N. (2024). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. IEEE Access.

Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. Ieee Access, 10, 99129-99149.

Mir, A. A. (2024). Adaptive Fraud Detection Systems: Real-Time Learning from Credit Card Transaction Data. Advances in Computer Sciences, 7(1).

Mitchell, C. (2023). Identity Theft Prediction Model using Historical Data and Supervised Machine Learning: Design Science Research Study (Doctoral dissertation, Colorado Technical University).

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

Mohbey, K. K., Khan, M. Z., & Indian, A. (2022). Credit card fraud prediction using XGBoost: an ensemble learning approach. International Journal of Information Retrieval Research (IJIRR), 12(2), 1-17.

Muaz, A., Jayabalan, M., & Thiruchelvam, V. (2020). A comparison of data sampling techniques for credit card fraud detection. International Journal of Advanced Computer Science and Applications, 11(6).

Mukherjee, M., & Khushi, M. (2021). SMOTE-ENC: A novel SMOTE-based method to generate synthetic data for nominal and continuous features. Applied System Innovation, 4(1), 18.

Mwangi, E. (2024). Employing AI/ML to Determine and Mitigate Fraud in the Insurance Industry. Available at SSRN 4907329.

Naidu, G., Zuva, T., & Sibanda, E. M. (2023, April). A review of evaluation metrics in MLalgorithms. In Computer science on-line conference (pp. 15-25). Cham: Springer International Publishing.

Ndungu, G. M. (2021). Detecting zero-day attacks using Recurrent Neural Network (Doctoral dissertation, Strathmore University).

Ngo, G., Beard, R., & Chandra, R. (2022). Evolutionary bagging for ensemble learning. Neurocomputing, 510, 1-14.

Njoku, D. O., Iwuchukwu, V. C., Jibiri, J. E., Ikwuazom, C. T., Ofoegbu, C. I., & Nwokoma, F. O. (2024). MLapproach for fraud detection system in financial institution: A web base application. Machine Learning, 20(4), 01-12.

Noviandy, T. R., Idroes, G. M., Maulana, A., Hardi, I., Ringga, E. S., & Idroes, R. (2023). Credit card fraud detection for contemporary financial management using XGBoost-driven machine learning and data augmentation techniques. Indatu Journal of Management and Accounting, 1(1), 29-35.

Nuthalapati, A. (2023). Smart fraud detection leveraging MLfor credit card security. Educational Administration: Theory and Practice, 29(2), 433-443.

Nweze, M., Avickson, E. K., & Ekechukwu (2024), G. The Role of AI and MLin Fraud Detection: Enhancing Risk Management in Corporate Finance.

Obi, J. C. (2023). A comparative study of several classification metrics and their performances on data. World Journal of Advanced Engineering Technology and Sciences, 8(1), 308-314.

Okenwa, C. D., David, O. D., Orelaja, A., & Akinwande, O. T. (2024). Exploring the Role of Explainable AI in Compliance Models for Fraud Prevention. International Journal of Research and Scientific Innovation, 13(5), 232-239.

Olushola, A., & Mart, J. (2024). Fraud Detection using Machine Learning. ScienceOpen Preprints.

Pandey, J. (2019). Deductive approach to content analysis. In Qualitative techniques for workplace data analysis (pp. 145-169). IGI Global.

Parkinson de Castro, E. (2020). An examination of the smote and other smote-based techniques that use synthetic data to oversample the minority class in the context of credit-card fraud classification.

Patel, H. I., & Patel, D. (2024). Exploratory Data Analysis and Feature Selection for Predictive Modeling of Student Academic Performance Using a Proposed Dataset. Int. J. Eng. Trends Technol, 72, 131-143.

Patel, K. (2023). Credit card analytics: a review of fraud detection and risk assessment techniques. International Journal of Computer Trends and Technology, 71(10), 69-79.

Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., & Turini, F. (2019, July). Meaningful explanations of black box AI decision systems. In Proceedings of the AAAI conference on artificial intelligence (Vol. 33, No. 01, pp. 9780-9784).

Plakandaras, V., Gogas, P., Papadimitriou, T., & Tsamardinos, I. (2022). Credit card fraud detection with automated MLsystems. Applied Artificial Intelligence, 36(1), 2086354.

Poddar, H. (2024). From neurons to networks: Unravelling the secrets of artificial neural networks and perceptrons. In Deep Learning in Engineering, Energy and Finance (pp. 25-79). CRC Press.

Porwal, U., & Mukund, S. (2019, August). Credit card fraud detection in e-commerce. In 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) (pp. 280-287). IEEE.

Priscilla, C. V., & Prabha, D. P. (2021). A two-phase feature selection technique using mutual information and XGB-RFE for credit card fraud detection. Int. J. Adv. Technol. Eng. Explor, 8(85), 1656-1668.

Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. Wiley Interdisciplinary Reviews: data mining and knowledge discovery, 9(3), e1301.

Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. Scientific Reports, 14(1), 6086.

Rajeev, H., & Devi, U. (2022). Detection of credit card fraud using isolation forest algorithm. In Pervasive Computing and Social Networking: Proceedings of ICPCSN 2021 (pp. 23-34). Springer Singapore.

Rasistia, W. P., & Sayyidah, M. J. (2021). PERCEIVED SECURITY AND TRUST IN ELECTRONIC PAYMENT SYSTEMS_ HOW THEY AFFECT THE DECISION TO USE EPS DURING THE COVID-19 PANDEMIC. Jurnal Manajemen Bisnis, 12(2), 236-247.

Razikin, K., & Widodo, A. (2021). General cybersecurity maturity assessment model: Best practice to achieve payment card Industry-Data security standard (PCI-DSS) compliance. CommIT (Communication and Information Technology) Journal, 15(2), 91-104.

Rezapour, M. (2019). Anomaly detection using unsupervised methods: credit card fraud case study. International Journal of Advanced Computer Science and Applications, 10(11).

Richardson, E., Trevizani, R., Greenbaum, J. A., Carter, H., Nielsen, M., & Peters, B. (2023). The ROC-AUC accurately assesses imbalanced datasets. Available at SSRN 4655233.

Salman, H. M. (2024). Identity Theft in the Banking System. In Online Identity-An Essential Guide. IntechOpen.

Shams, S. R., Sobhan, A., & Vrontis, D. (2021). Detection of financial fraud risk: implications for financial stability. Journal of Operational Risk.

Sharma, S., & Sehgal, V. (2024). Credit Card Fraud Detection.

Singla, J., Bashir, A. K., Nam, Y., Hasan, N. U., & Tariq, U. (2021). Handling class imbalance in online transaction fraud detection. Computers, Materials and Continua, 70(2), 2861-2877.

Tan, J., Yang, J., Wu, S., Chen, G., & Zhao, J. (2021). A critical look at the current train/test split in machine learning. arXiv preprint arXiv:2106.04525.

Tarlin, S. (2021). The future of cash. Federal Reserve Bank of Philadelphia Discussion Paper, 20-03.

Technology and Operations Management. (2018). PayPal's Use of Machine Learning to Enhance Fraud Detection (and more) - Technology and Operations Management. [online] Available at: https://d3.harvard.edu/platform-rctom/submission/paypals-use-of-machine-learning-to-enhance-fraud-detection-and-more/ [Accessed 20 Jun. 2025].

Tiwari, M., Sharma, V., & Bala, D. (2022). Credit card fraud detection. Journal of Algebraic Statistics, 13(2), 1778-1789.

Trivedi, N. K., Simaiya, S., Lilhore, U. K., & Sharma, S. K. (2020). An efficient credit card fraud detection model based on MLmethods. International Journal of Advanced Science and Technology, 29(5), 3414-3424.

Udeh, E. O., Amajuoyi, P., Adeusi, K. B., & Scott, A. O. (2024). The role of big data in detecting and preventing financial fraud in digital transactions. World Journal of Advanced Research and Reviews, 22(2), 1746-1760.

Vagadia, B. (2020). Data integrity, control and tokenization. In Digital Disruption: Implications and opportunities for Economies, Society, Policy Makers and Business Leaders (pp. 107-176). Cham: Springer International Publishing.

Van der Cruijsen, C., De Haan, J., & Roerink, R. (2023). Trust in financial institutions: A survey. Journal of economic surveys, 37(4), 1214-1254.

Vaquero, P. R. (2023). LITERATURE REVIEW OF CREDIT CARD FRAUD DETECTION WITH MACHINE LEARNING.

Voican, O. (2021). Credit Card Fraud Detection using Deep Learning Techniques. Informatica Economica, 25(1).

Vorobyev, I., & Krivitskaya, A. (2022). Reducing false positives in bank anti-fraud systems based on rule induction in distributed tree-based models. Computers & Security, 120, 102786.

Wang, T., & Zhao, Y. (2022, January). Credit card fraud detection using logistic regression. In 2022 International Conference on Big Data, Information and Computer Network (BDICN) (pp. 301-305). IEEE.

Zhang, C., Wang, Q., Liu, T., Lu, X., Hong, J., Han, B., & Gong, C. (2021, October). Fraud detection under multi-sourced extremely noisy annotations. In Proceedings of the 30th ACM international conference on information & knowledge management (pp. 2497-2506).

Zhou, Z., & Hooker, G. (2021). Unbiased measurement of feature importance in tree-based methods. ACM Transactions on Knowledge Discovery from Data (TKDD), 15(2), 1-21.

Zou, J., Zhang, J., & Jiang, P. (2019). Credit card fraud detection using autoencoder neural network. arXiv preprint arXiv:1908.11553.