Fraud Detection in Financial Transactions: Using Machine

Learning to Identify Fraudulent Transactions and

Prevent Financial Losses

*A dissertation submitted in fulfilment of the requirements*

*For the degree of Master of Science*

*By*

*Jones Kwame Osei*

*S4112106*

*A picture containing clipart, crest, illustration

Description automatically generated*

SCHOOL OF COMPUTING AND ENGINEERING

UNIVERSITY OF GLOUCESTERSHIRE

*Supervised by*

Dr. Bhupesh Mishra

June 2023

Contents

[1 List of Figures iv](#_Toc137185077)

[2 List of Tables v](#_Toc137185078)

[3 Abstract vi](#_Toc137185079)

[3.1 Keywords vi](#_Toc137185080)

[4 Chapter 1 – Introduction 1](#_Toc137185081)

[4.1 Background and Motivation 1](#_Toc137185082)

[4.2 Problem Statement 2](#_Toc137185083)

[4.3 Aims and Objectives 3](#_Toc137185084)

[4.4 Research Questions and Hypotheses 4](#_Toc137185085)

[4.5 Significance of this Study 4](#_Toc137185086)

[4.6 Scope and Limitations 4](#_Toc137185087)

[4.7 Ethical Issue 5](#_Toc137185088)

[4.8 Conclusion 5](#_Toc137185089)

[4.9 Outline of this study 5](#_Toc137185090)

[5 Chapter 2 – Literature Review 7](#_Toc137185091)

[5.1 Overview of Fraud Detection in Financial Transactions. 7](#_Toc137185092)

[5.2 Problems Confronting the Detection of Credit Card 7](#_Toc137185093)

[5.3 The Role of Machine Learning in the Financial Sector and Related Work 8](#_Toc137185094)

[5.4 Machine Learning-based Fraud Detection: 9](#_Toc137185095)

[5.5 Feature Engineering for Fraud Detection: 11](#_Toc137185096)

[5.6 Comparative Studies on Evaluation Metrics 12](#_Toc137185097)

[5.7 Explainable AI in Fraud Detection 13](#_Toc137185098)

[5.8 Conclusion and Study Gap 14](#_Toc137185099)

[6 Chapter 3: Research Methodology 15](#_Toc137185100)

[6.1 Understanding Business Problem 15](#_Toc137185101)

[6.2 Acquisition of Data 16](#_Toc137185102)

[6.2.1 Source of Data and Descriptions 16](#_Toc137185103)

[6.3 Data Inspection and Cleaning 17](#_Toc137185104)

[6.3.1 Basic Exploratory Data Analysis 17](#_Toc137185105)

[6.3.2 Data Preprocessing 17](#_Toc137185106)

[6.3.3 Statistical Analysis 18](#_Toc137185107)

[6.3.4 Dealing with Outliers with Interquartile Range (IQR) 21](#_Toc137185108)

[6.4 Data Imbalanced and Strategies for Handling it 23](#_Toc137185109)

[6.5 Machine Learning Modelling 25](#_Toc137185110)

[6.5.1 Overview of the Models Used 25](#_Toc137185111)

[6.5.2 Selection of Model and Justification 26](#_Toc137185112)

[6.5.3 Splitting the Data and Scaling 26](#_Toc137185113)

[6.5.4 Resampling techniques 27](#_Toc137185114)

[6.5.5 Hyperparameter Tuning Techniques 31](#_Toc137185115)

[6.5.6 Cross Validation 32](#_Toc137185116)

[6.5.7 Performance Evaluation Metrics 33](#_Toc137185117)

[7 Chapter 4: Results and Discussions 36](#_Toc137185118)

[7.1 Main Model with Default Parameters 36](#_Toc137185119)

[7.2 Optimised Main Model 38](#_Toc137185120)

[7.3 Oversample Model 40](#_Toc137185121)

[7.4 Optimised Oversample Model 42](#_Toc137185122)

[7.5 Undersample Model 43](#_Toc137185123)

[7.6 Interpreting Model Results (Explainable AI) 45](#_Toc137185124)

[7.7 Conclusion and Discission 47](#_Toc137185125)

[7.7.1 Critical Analysis of This Study 48](#_Toc137185126)

[8 Reference List 50](#_Toc137185127)

# List of Figures

[Figure 6‑1*. Project Process Flow* 15](#_Toc137181651)

[Figure 6‑2*. Visualising missing values* 18](#_Toc137181652)

[*Figure 6‑3. Distribution of Amount* 19](#_Toc137181653)

[*Figure 6‑4. Outliers Detection with Histogram* 19](#_Toc137181654)

[*Figure 6‑5. Detecting Outliers with Boxplot* 20](#_Toc137181655)

[*Figure 6‑6. Dataset structure after removing Outliers* 22](#_Toc137181656)

[*Figure 6‑7. Dataset Class Count* 24](#_Toc137181657)

[*Figure 6‑8. Proportion of Class Distribution* 24](#_Toc137181658)

[*Figure 6‑9. Correlation Matrix Before Reduction* 28](#_Toc137181659)

[*Figure 6‑10. Undersampling by Dimensionality Reduction* 29](#_Toc137181660)

[*Figure 6‑11. Correlation after Dimension Reduction* 30](#_Toc137181661)

[*Figure 7‑1. AUC Score Plot from Main Model* 37](#_Toc137181662)

[*Figure 7‑2. Stack Classifier for Mian Model* 37](#_Toc137181663)

[*Figure 7‑3 AUC Score Plot for Optimised Main Model* 38](#_Toc137181664)

[*Figure 7‑4. Stack Classifier for Optimised Mian Model* 39](#_Toc137181665)

[*Figure 7‑5. AUC Plot for Oversample Model.* 40](#_Toc137181666)

[*Figure 7‑6. AUC Plot for Oversample Voting Model* 41](#_Toc137181667)

[*Figure 7‑7. AUC Plot for Oversample Stacking Model* 41](#_Toc137181668)

[*Figure 7‑8.AUC Plot for Optimised Oversample Model* 42](#_Toc137181669)

[*Figure 7‑9. AUC Plot for Optimised Oversample Voting Classifier* 43](#_Toc137181670)

[*Figure 7‑10. AUC Plot of Undersample Data* 44](#_Toc137181671)

[*Figure 7‑11. AUC Plot of Undersample Stacking Classifier* 45](#_Toc137181672)

[*Figure 7‑12. Explainable AI - Features that Model Flags as Fraud* 46](#_Toc137181673)

[*Figure 7‑13. Explainable AI - Features that Model Flags as No Fraud* 46](#_Toc137181674)

[*Figure 7‑14. Random Forest Feature Importance* 47](#_Toc137181675)

# List of Tables

[*Table 1. Distribution of Classes in the dataset* 25](#_Toc137181684)

[*Table 2. Dimension of Dataset after Applying Rus* 28](#_Toc137181685)

[*Table 3. Dimension of Dataset after Applying NearMiss Version-3* 31](#_Toc137181686)

[*Table 4. Traditional Confusion Matrix* 33](#_Toc137181687)

[*Table 5. Summary of Performance Metrics* 35](#_Toc137181688)

[*Table 6. Results from Main Model* 36](#_Toc137181689)

[*Table 7. Results from Optimised Main Model* 38](#_Toc137181690)

[*Table 8. Results from Oversample Model* 40](#_Toc137181691)

[*Table 9. Results from Optimised Oversample Model* 42](#_Toc137181692)

[*Table 10. Results from Undersample Model* 43](#_Toc137181693)

# Abstract

Credit card fraud has become a pressing subject matter in recent times, leading to substantial financial losses. This social menace is increasing because of the prevalence of electronic payment systems (EPS). Techniques of machine learning have been widely adopted to identify fraudulent transactions in financial ecosystems and ensure security. As fraudsters have become more vicious in the search for credit card transactions, the financial sector has to make more stringent decisions to curb or eradicate the root of this problem, which must be a data-driven solution. Hence, in this dissertation project, we focus on building, optimising and comparing the performance of diverse machine learning classifiers as well as employing different data balancing strategies for detecting and preventing fraud. Seven main models including random forest, decision tree, extreme gradient boosting (xgboost), extreme gradient boosting random forest, logistic regression, k-nearest neighbour and stochastic gradient were built. Our results revealed that the optimised models built from original and oversample data performed better than their actual models. Generally, classifiers on undersample data had the worst performance using precision, recall, precision-recall and receiver operating characteristics curve (ROC) and area under curve (AUC) as the performance metrics. Overall, our findings suggest that machine learning models can effectively be used to detect trends, patterns and unusual behaviour in financial transactions than the traditional rule-based system. It further stipulates that for efficient results, optimisation techniques and ensemble algorithms can further improve their performance.

## Keywords

Electronic payment systems (EPS); Fraudulent transactions; Genuine transactions; Financial ecosystems; Data-driven solution; Machine Learning (ML); Classification algorithms; Optimisation; Hyperparameters; Decision Tree; Random Forest; Logistic Regression; Stochastic Gradient (SGD); K-Nearest Neighbour; extreme gradient boosting (xgboost).

# Chapter 1 – Introduction

Fraud detection in banking, electronic commerce (e-commerce) and financial transactions is an important aspect of considering business sustainability and growth and received the needed industrial attention with the predominance of digital transactions. The implications of credit card fraud are capable of causing serious financial losses, endangering not only banks but also the clients who fall victim to these fraudulent activities. It was reported that 2.6 million people experienced emotional distress due to victimisation by debit or credit card fraud in the USA (Burnes et al., 2020). Therefore, the need to discover and prevent such fraudulent activities is of utmost importance. The traditional methods of detecting and preventing fraud activities have been obsolete and, hence, machine learning algorithms are progressively being used to identify fraudulent transactions from genuine ones and help prevent financial loss.

## Background and Motivation

In general, fraud is a serious problem that imposes great danger on various sectors and industries, specifically the banking and finance sectors. Fraudulent transactions within these sectors can trigger significant financial losses and potentially damage the reputation and trust of the stakeholders and customers. Right from the emergence of the internet, credit card fraud has always been inevitable, and the advancement of technology has also seen an increase in this criminal activity. Traditionally, before the application of the first ever internet banking by Wells Fargo Bank, United States of America in 1995 banks used to just provide in-person customer service (Tudeal, 2022). Internet banking was made possible because a group of New York's major banks including Manufacturers Hanover, Chase Manhattan, Chemical Bank and Citibank failed to provide home banking in 1981 through the videotex system. However, videotex became a success in France after its launch in 1983 by the Crédit Commercial de France (CFF) Bank (now a subsidiary of HSBC) and ultimately reached 19% of the market share in 1991. This system continued to serve customers all over the world delivery information through text in pages to the early 2010s (McKinlay, 1982; Philip, 2018). Notwithstanding its success, the success of Internet banking and its user interface has gained more users and subsequently has become a popular banking application. Given this, there have been recent advancements in digital technologies, specifically for cash transactions, changing how people manage their daily activities. Most transactions are now done through digital platforms, especially using credit (contactless) cards (Yash, 2018; Windasari et al., 2022).

With the proliferation of banking online, the use of credit cards increased tremendously in the last two decades. There has also been a wide use of services like online mobile banking, google pay and mobile money transfers making it simple for customers to make transactions and interact with their accounts at the comfort of wherever they are (Hodjat, 2019; Singh et al., 2020). Despite the technological advantage of these digital applications, there have been adverse implications of abuse and misuse of these digital transformations such as credit card fraud. This form of fraud is the unlawful use of an electronic payment card belonging to another person to make a purchase and/or transfer either physically or on a digital platform without the owner’s consent (Carcillo *et al.*, 2021a). It is easier now for criminals to commit card fraud than before because of the increased access to the dark web as well as technological methods and tools such as spoofing, virtual proxy networks (VPNs) and connecting to socks proxy (changing their location to card holder’s postcode and addresses) allows these fraudsters to have details of credit customers (Thorat et al., 2020; Jung et al, 2022). With these in the hands of criminals, credit card fraud activities are made simple and without any fear of getting apprehended, a huge sum of money can effortlessly be stolen without the rightful owner knowing. That is, these criminal perpetrators have achieved great success in making every fraudulent transaction look genuine which makes the detection of fraud a challenging task (Dornadula et al., 2019).

## Problem Statement

The global cost of fraud through electronic commerce in 2020 was estimated at $42 billion (PwC, 2022). Again, the 2020 internet crime report including data from 791,790 complaints revealed that credit card fraud had the highest(IC3, 2021) complaints and an estimated $4.2 billion were lost by organisations and individuals (IC3, 2021) in the United States of America alone. Credit card fraud contributed to a loss of £524.5 million in 2021 (UK Finance, 2022) in the United Kingdom. The Bank of Ghana (BoG)’s annual fraud report recorded an increase of 144% in 2021 representing an estimated loss of GH¢61 million ($5 million) as compared to the loss of GH¢25 million ($2.1 million) in 2020 (BoG, 2022). Debit or/and credit card fraud significantly accountant for the increase.

All payment and transaction methods have suffered a constant rise in fraud over the years and transactions made through digital platforms witnessing the highest (Chang *et al.*, 2022). The fast growth of the electronic commerce industry where people purchase goods and services and make payments online using debit or credit has led to the increase of fraud exponentially. Electronic commerce is essential to the sustainability of today’s businesses and industries, however, its prevalence especially during and after covid-19 pandemic has dramatically contributed to the overwhelming escalation in online payment fraud(Alawida et al., 2022; Mienye et al, 2023). Digital transactions offer several advantages to customers, including a quicker buying process, the flexibility of delivery at customers' doorsteps, and product reviews for comparison despite the risks of fraud. Therefore, detecting and preventing fraud is a crucial task for ensuring security and efficacy in the financial industry.

Attacks on credit cards are difficult to identify. To discover financial fraud, machine learning is used to help mitigate this ever-growing menace. Machine learning techniques can be used to discover fraudulent transactions. It can also efficiently handle big data, adaptable to dynamic environments and provide insights into the root causes and mechanisms of fraud. Several artificial intelligence algorithms have been built to empower the banking and finance industry to uncover fraud, however, applying these algorithms for fraud detection also comes with some challenges and still needs room for improvement. A major challenge in fraud detection is the problem of a biased dataset. A dataset is biased when the number of legal transactions is significantly greater than the number of fraudulent transactions. This affects the accuracy of the classifiers and can potentially lead to false negatives or positives as the models favours the majority class and ignore the minority class. Notwithstanding, data balancing techniques can mitigate this problem. Data balancing technique is the process of normalising the dissemination of the classes within data, either by decreasing the number of transactions classed as non-fraud (undersampling) or by increasing fraud occurrences (oversampling).

## Aims and Objectives

The study mainly aims at effectively applying and making comparisons of several machine learning models to ascertain the model with the highest efficacy for fraud detection in financial transactions, and exploring the effects of different data balancing techniques. The specific objectives include collecting, gathering and processing credit card transactions dataset with labels showing whether they are fraudulent or not. This research also seeks to implement and assess different machine learning techniques such as the original dataset. This study further seeks to harness different techniques to balance data such as oversampling, dimensionality reduction and NearMiss on the original dataset and make comparisons on their model performance. Furthermore, this study aims to identify the best metric for measuring the performance of models for the imbalance dataset and to discuss the findings and implications of the research, as well as its limitations.

## Research Questions and Hypotheses

This dissertation seeks to answer these research questions:

RQ1: What patterns and features are associated with fraudulent transactions and are machine learning algorithms able to accurately detect transactions classed as fraud in unbalanced data?

RQ2: Do different data balancing techniques influence how machine learning models perform?

RQ3: In what ways can models designed be optimised to improve their accuracy in making fraud predictions?

## Significance of this Study

This research has numerous relevant implications. Firstly, it will contribute to the body of knowledge on fraud detection in financial transactions. Secondly, the outcomes of this research will provide insights into effective and efficient ways of leveraging machine learning techniques and algorithms to detect fraud. Finally, this study will help the financial sector and other organisations to adjust and improve their fraud detection capabilities and prevent financial loss.

## Scope and Limitations

This study mainly focuses on utilising machine learning algorithms for the detection of fraud behaviours. The study carefully considers a wide range of machine learning classifiers, including decision trees, logistic regression, and k-nearest neighbours, as well as different sampling techniques such as reduction techniques to achieve data balance. Notwithstanding the study is limited to a single dataset and may not be generalisable to other datasets. Moreover, the use of an available publicly transactional dataset from Kaggle (Dal Pozzolo et al., 2015), may not replicate the real-world situation of credit card fraud detection in the banking and finance industry. Again, the use of these machine learning classifiers, the optimisation techniques, the techniques to balance the dataset and the metrics measuring the performance of the models may not cover all possible methods or combinations for fraud detection in financial transactions. Furthermore, the lack of domain expert feedback on credit card fraud in transacting business may affect the generalisability of the results. Lastly, the reliability of the models may be affected due to the possibility of overfitting or underfitting an imbalanced dataset.

## Ethical Issue

To conduct this study responsibly, all classifiers used should be accurate so that customers, stakeholders and financial institutions are not put into high risk (Caroline Cynthia et al. 2021). The handling and use of data will be adhered to and respected under all accepted ethical guidelines. A thorough reference will be provided for all existing work and knowledge reviewed and research guidance from the university will be followed accordingly. This study is my work in all sincerity, and it adheres to all referencing guidelines.

## Conclusion

The utilisation of machine learning algorithms in detecting fraudulent transactions has proven to be feasible. This dissertation aims to unearth the appropriate machine learning classifiers and techniques for fraud detection in financial transactions. It also adds to the body of domain knowledge on credit fraud detection and offers a broad spectrum of methodologies to handle financial fraud helping monetary institutions and other organisations to improve credit card fraud detection capabilities and ultimately prevent financial loss.

## Outline of this study

The remaining of the project is organised as follows:

* Chapter 2 reviews the relevant literature on fraud, credit card and digital transactions, card fraud, detecting credit card fraud, machine learning algorithms, and data balancing techniques as well as identifying research gaps and contributions.
* Chapter 3 describes the research methodology, design and approach, including data collection and description, data reprocessing and analysis, and machine learning model selection.
* Chapter 4 discusses the results and findings of this study.

# Chapter 2 – Literature Review

## Overview of Fraud Detection in Financial Transactions.

In the banking and finance industry events such as fraud can be described as a business threat to institutions, individuals, and an entire economy. In an attempt to expose and prevent fraud, it has become a critical area of investigation and research. Literature on fraud detection highlights the nature of fraudulent transactions and the confronting issues with detecting them. Many strategies have been proposed to either prevent or significantly reduced the occurrence of fraud in earlier studies. To detect fraud, researchers have suggested approaching this problem by utilising supervised, unsupervised or the combination (Kim *et al.*, 2019) of the two into a hybrid approach, making it a necessity to learn the current technologies related to fraud in financial transactions and to have a comprehensive idea of the modus operandi of credit card fraud as well as the types.

Fraudulent activities come in the form of the following, such as credit card fraud, identity theft, impersonation, and insider trading. As these activities typically involve complicated techniques and evolving patterns constantly, they are difficult to detect by traditional rule-based systems. A growing need for advanced methods and technologies to combat fraud has arisen as a result. Traditionally, fraud detection has been carried out via rule-based systems, which utilize predefined thresholds and rules to identify suspicious activity. They can flag potential fraud cases, but they produce many false positives and struggle to adapt to changing fraud patterns. As time elapsed the patterns of fraud also evolved sprouting out new approaches criminals use making it an interesting area for researchers to study, be abreast with time and meet the ever-growing demands of consumers. Describing multiple machine learning algorithms, machine learning classifiers, systems for detecting financial fraud and how to approach this problem will be the focus of the remaining part of this section.

## Problems Confronting the Detection of Credit Card

Analysis of the problems identified in the review has been done to develop a more efficient machine learning model in the future. In recent times, researchers in this domain have highlighted a lot of industry problems confronting credit card issuers and users with the analysis of several detection models. Inadequate quality data (Dal Pozzolo et al., 2014; Munappy et al., 2022) is one issue confronting the detection of card fraud. From the reviews, the majority of studies relied hugely on historical data, and this is similar to this data source. Another caveat about data acquisition is the hiding of certain features in most of this historic public dataset for confidentiality and security reasons. This is a hindrance to study patterns and characteristics which are mostly related to transactional fraud. Even the readily available historic data present another growing concern in the domain, the presence of imbalanced classes. Many of the datasets reviewed are biased toward transactions classified as non-fraudulent. These transactions are the majority class making class distribution imbalanced and this occurs because (Mohammed et al., 2018; Makki et al., 2019). Throughout the review, most of the studies propose one classifier or the other to be effective for detecting transactional fraud. However, when one study affirms an algorithm to be the most effective, another study refutes it creating the problem of choosing a suitable model (Sailusha *et al.*, 2020) to tackle the fraud detection problem. For instance, whereas Sohony et al, (2018) propose random forest to be the best algorithm to solve the problems of fraud, Taneja et al, (2019) recommend support vector machine is mostly credible as compared to random forest.

## The Role of Machine Learning in the Financial Sector and Related Work

Detecting fraud has been a critical challenge within the field of financial transactions and the classification of credit card fraud is mostly considered as a binary classification problem. That is, debit or credit card transaction is either seen as genuine (negative class) or fraudulent transaction (positive class) (Awoyemi et al., 2017). Traditional rule-based approaches have long been used for detecting fraudulent transactions. However, these manual methods often struggle to keep up with the advanced nature of fraud patterns and are unable to capture complex relationships in the data. This setback has led to the adoption of machine learning and artificial intelligence procedures in fraud detection. Machine learning approaches offer the potential to analyse large volumes of transactional data and identify subtle trends and patterns related to fraudulent behaviours. Researchers have explored the application of various models such as neural networks, random forest and support vector machine for credit card fraud detection. These models learn from historical data and adapt to new dynamics in fraud tactics, making them well-suited for detecting both known and emerging fraud patterns.

## Machine Learning-based Fraud Detection:

Several studies have shown the efficacy of machine learning classifiers in fraudulent behaviours in financial transactions. Techniques of machine learning have gained traction in detecting fraud in recent years because of their capability to learn and adapt to big data. In these techniques, complex patterns and anomalies are automatically identified as indicators of fraudulent behaviour. Some supervised learning has been successfully applied to discover fraud from labelled data such as support vector machines, random forest, and neural networks. It has been demonstrated that unsupervised learning techniques including clustering and anomaly detection, may assist in identifying fraudulent transactions without labelling the data. Outliers and patterns that deviate significantly from the norm can be identified using these methods. Machine learning algorithms offer a powerful approach to fraud detection by leveraging the vast amounts of transactional data available. These algorithms can learn from historical data to detect fraudulent patterns and adapt to contemporary fraud schemes in real-time. Ensemble methods, such as gradient boosting and stacking are also used to improve the performance of fraud detection classifiers. These methods combine multiple models to make more accurate predictions by aggregating their outputs.

Credit card fraud detection, for instance, was described as a sequence classification task by Jurgovsky et al., as well as the application of Long Short-Term Memory (LSTM) networks in incorporating transaction series to detect fraud. This research also compared Long Short-Term Memory Networks to a baseline random forest and the former gave an improved accuracy in detecting offline transactions in which the cardholder is physically present at a sales point (Jurgovsky *et al.*, 2018). A neural network-based system called APATE was Also employed by Van Vlasselaer et al., to find fraud patterns in credit card transactions. Through the use of the network of credit card holders and merchants, they combined the features of incoming transactions with the history of customer spending bahaviour. Consequently, both network and intrinsic characteristics intertwine (Van Vlasselaer *et al.*, 2015).

Alternatively, Xuan et al. highlighted that credit card fraud occurs frequently and leads to enormous financial losses. They further stressed that credit card information can be stolen through Trojans or phishing. It is therefore imperative to implement an effective fraud detection process. By doing so, fraud can be detected on time when an online buyer uses a stolen credit or debit card. One way to accomplish this is by making use of historic data, including non-fraudulent transactions and fraudulent ones. In this case, we are trying to obtain features that identify normal and fraudulent behaviours by implementing machine learning techniques. Once these features have been identified, they can be used to determine whether or not, a transaction is genuine. For training behavioural features of normal and abnormal transactions, random forests were applied to two types. A random-tree-based random forest classifier and a cart-based random forest classifier were made, and their performances in detecting credit card fraud were compared. They concluded that although the models gave good results, the algorithms need improvement as the classifiers showed different important features with the same equal weight (Xuan et al., 2018a).

Furthermore, the work of Dejan et al., highlighted the deployment of supervised models to find patterns of electronic payments fraud. They also performed feature selection and used different supervised models in their experiment. In this experiment, logistic regression, random forest, naive bayes, and multilayer perceptron were applied. The outcome indicates that each of the models used for credit card fraud detection had high prediction accuracy. The proposed models can be deployed for the detection of other irregularities (Varmedja *et al.*, 2019).

Whilst many of these research works leveraged supervised machine learning techniques, Fabrizio et al., identified the weaknesses in labelling target values and utilising customer transactional historic data to predict future transactional behaviours. Hence, they proposed a combination of both supervised and unsupervised techniques. They recognised that many supervised learning techniques are employed in detecting fraud based on previous transactions. However, fraudsters have the ability in inventing new modus operandi, which makes the task challenging. Therefore, unsupervised learning techniques can be useful in detecting fraud in this context. The purpose of their research was to present a hybrid fraud detection method combining both supervised and unsupervised strategies to increase fraud detection accuracy. An annotated, real credit card fraud dataset is used to examine the impact of outlier scores computed at various degrees of granularity. Based on their experimental results, this combination was efficient and improved the detection accuracy (Carcillo et al., 2021).

## Feature Engineering for Fraud Detection:

Feature engineering plays a crucial role in machine learning-based fraud detection. Relevant features, such as transaction amount, location, time, and customer behaviour, are extracted and used as input variables for the models. Feature selection techniques, such as information gain, correlation analysis, and recursive feature elimination, help identify the most informative features for fraud detection. In developing robust fraud detection models, feature engineering is crucial. A lot of attention is paid to selecting relevant features and transforming raw data into meaningful representations. In fraud detection, data obtained and correlation analysis are used to identify the most discriminative features. Scaling, normalization, and categorical encoding are all methods for transforming features to ensure compatibility with machine learning.

To address the class imbalance and improve model performance, researchers have employed sampling techniques such as oversampling, undersampling, and synthetic data generation. For instance, Khemakhem et al. (2018) used the synthetic minority oversampling technique (SMOTE) to solve class imbalance problems. They opined that because of imbalanced datasets, support vector machines (SVMs) are less effective at giving good prediction accuracy scores. In their study, a support vector machine model with more kernels was made and two data resampling alternatives were proposed: random oversampling (ROS) and synthetic minority oversampling (SMOTE). Their study explored the possibility of resampling data to derive an accurate prediction rate under class imbalance constraints. The following criteria were used to assess the performance of the proposed technique: accuracy, G-mean, sensitivity specificity, receiver operating characteristic curve area under the curve (AUC), and error type I and error type II. A real, imbalanced credit card dataset of a Tunisian bank was used to generate a significant empirical result from support vector machine sampling strategies. These results showed a performance improvement, improving prediction accuracy.

In the same way, Sahu et al. 2020 emphasised the significance of adopting a robust approach to fraud detection within the banking industry. Their research highlighted the potential to prevent fraudulent transactions and save substantial financial resources annually. Detecting and mitigating fraudulent activities is imperative in every monetary transaction. The study focused on developing models to identify illegal credit card transactions by employing five different classifiers. It determined the most effective one for the given scenario. In addressing the challenges of data imbalance, the researchers employed two distinct techniques. Firstly, they increase the number of samples in the minority class. Secondly, they implemented a cost-based strategy where the error function incorporated weights for both classes. The findings of the study demonstrated that by incorporating weights, it was possible to effectively highlight instances of fraudulent transactions over normal transactions.

## Comparative Studies on Evaluation Metrics

Various fraud detection models are evaluated and measured through comparative studies. Researchers have comprehensively evaluated the performance of many algorithms. They have also looked into other methods such as rule-based and expert systems. The resulting models have been used to detect fraud with varying degrees of accuracy. Evaluation metrics used to measure fraud detection effectiveness include accuracy precision, recall, F1-scores, and areas under curves and receiver operating characteristic curves (AUC-ROC).

For instance, Bhattacharyya et al. (2011) conducted a comparative analysis of various models for detecting credit card fraud. The evaluation encompassed key performance metrics including accuracy, precision, specificity, F-measure, recall (sensitivity), and confusion matrix. The study revealed that ensemble models, specifically gradient boosting and random forests, outperformed individual models in overall performance. Bhattacharyya et al. concluded that ensemble models provided a more robust and accurate way to detect fraud. Furthermore, they suggested that combining multiple models should be considered when developing fraud detection systems.

Notwithstanding, the measurement or evaluation of algorithms is not constant as it varies or depends on many factors as elaborated by Mittal et al., (2019). They highlighted that traditional accuracy measures and confusion matrices are insufficient for capturing the true identification rate of fraud. This is possible because of how skewed the dataset is. Algorithms are usually biased towards the majority class when dealing with imbalanced data, leading to predictions that favour the majority class. Therefore, it becomes necessary to utilize alternative metrics such as the G-mean score, precision, recall, or receiver operating characteristic curve/area under curve. These modified metrics consider the skewed distribution of the classes and consider the costs associated with misclassifying different classes. This provides a more accurate assessment of fraud detection rates. It is essential to use metrics that strike a balance between detecting fraud instances from both the minority and majority classes.

These comparative studies provide insight into the strengths and limitations of different fraud detection models, as well as their metrics for assessing performance in real-world situations. Data scientists can benefit from these researchers in selecting appropriate algorithms and techniques that fit the fraud detection task at hand.

## Explainable AI in Fraud Detection

Obtaining trust and providing insights into decision-making requires the interpretability of fraud detection models. Complex models are black boxes that have drawn attention to explainable AI techniques. Explainable AI techniques focus on providing transparency and interpretability to machine learning models and systems. This helps to identify and explain the factors and decisions involved in the predictions made by the model. Such techniques can help increase user trust and acceptance of the model, as well as identify areas for improvement. By using these techniques, stakeholders will be able to understand and validate why fraud detection decisions are made, and they will be able to provide transparent explanations for these predictions (Miller, 2019).

Local Interpretable Model-Agnostic Explanations (LIME) is one of the many methods that have been applied in fraud detection to provide explanations for individual machine learning model predictions. The method approximates the behaviour of a complex model locally and identifies the most influential features contributing to the prediction. By highlighting relevant features, LIME enables users to understand why a particular transaction was classified as fraudulent or non-fraudulent. Demonstrating its effectiveness in explaining black-box models' predictions, Ribeiro et al. (2016) introduced the LIME framework. It has been effectively utilised in other domains as well as fraud detection (Wu et al., 2021). Stakeholders can gain insights into the features driving the predictions and increase trust in the model's decisions. This is done by integrating LIME into fraud detection models.

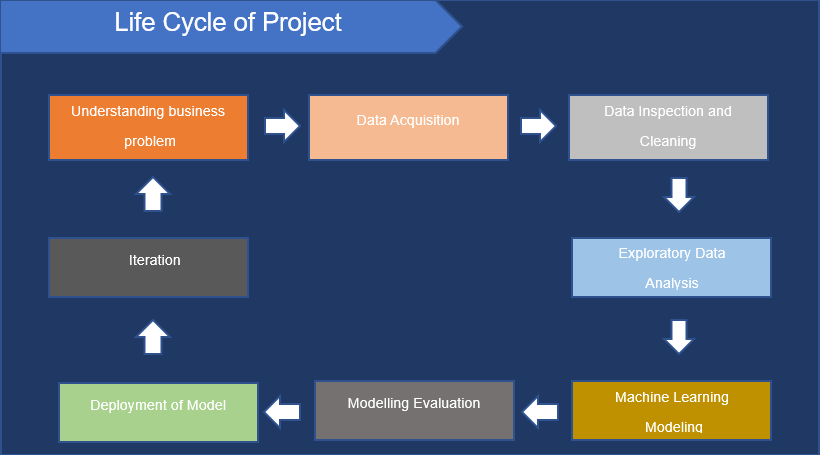
## Conclusion and Study Gap

Despite the efficacy of machine learning in the detection of fraud, challenges still exist. An uneven dataset poses a major hurdle. The rarity of fraud cases could potentially lead to biased models that prioritize accuracy on the majority class while overlooking the minority class. Addressing this imbalance problem, the reviewed literature failed to address the overlapping of data when preparing transaction datasets for credit card transactions. The reviewed studies also failed to address the scenarios when genuine transactions appear fraudulent, and the snail-paced data mining techniques fail to spot these inconsistencies because they take longer time to execute tasks when dealing with big data. In as much as the papers exposed inadequate data, they failed to highlight the need for real-time transactional data that will train models to be abreast with current trends in detecting fraud. Instead of focusing more on best techniques, researchers should be more concerned with model errors where legitimate transactions are flagged as fraudulent, and these should be the focal points of future studies.

In conclusion, the literature on fraud detection highlights the growing importance of advanced techniques, particularly machine learning, in identifying and preventing fraudulent activities in financial transactions. By leveraging the power of algorithms, feature engineering, and ensemble methods, researchers have made significant strides in improving fraud detection accuracy. However, further research is needed to address challenges related to imbalanced datasets, and real-time processing and to develop robust and adaptable models capable of detecting emerging fraud patterns.

# Chapter 3: Research Methodology

Figure 6‑1*. Project Process Flow*



This Chapter elaborates on the selection of the dataset for the research and also demonstrates a comprehensive account of diverse models and strategies utilised in creating the proposed credit card detection models.

## Understanding Business Problem

Fraud detection systems that deploy traditional rule-based approaches often struggle to keep up with the changing nature of fraudulent activity as technology advances (Zhou et al., 2018). In this context, algorithms provide a promising solution by combining data and pattern recognition to identify and flag suspicious transactions (Sadikin et al., 2020; Awotunde et al., 2021). Businesses, organisations and customers in a broad spectrum benefit from the adoption of machine learning in detecting credit card transactional fraud. A machine learning model can examine a large amount of transactional data to detect complex patterns and behaviours that are indicative of fraud. It enables businesses to recognise fraudulent transactions quickly, reducing financial losses. As new fraud patterns become apparent, machine learning algorithms can evolve, improving detection capabilities and staying ahead of fraudsters.

Automating fraud detection processes with machine learning-based systems can help businesses especially financial institutions reduce manual labour of fraud prevention and improve operational efficiency (Kim et al., 2019). Early and accurate detection of fraud helps businesses safeguard their financial interests and protect their customers' assets. Additionally, incorporating machine learning techniques to discover fraudulent behaviours aligns with a broader trend of leveraging data-driven insights to drive business outcomes. Machine learning algorithms have the potential to assist businesses to find useful insights, customer behaviour patterns, and risk factors. It is possible to use these insights to inform strategic decision-making, risk management strategies, and more targeted fraud prevention techniques (Aziz et al., 2019). Using machine learning to detect fraud is essential from a business perspective to understanding how it happens. By deploying sophisticated machine learning models businesses can boost their fraud detection capabilities, reduce losses, and ensure customer trust. By embracing machine learning-based fraud detection, businesses not only strengthen their defenses against fraudulent activities but also enhance their risk mitigation (Leo et al. 2019) strategy by leveraging data-driven insights.

## Acquisition of Data

### Source of Data and Descriptions

The data for this dissertation has been drawn from the Credit Card Fraud Detection dataset, which is readily available through the Kaggle platform. It is a real credit card transaction by card users in Europe. This dataset contains a large number of credit card transactions, including both fraudulent transactions as well as non-fraudulent transactions. Various transactional characteristics are captured in the dataset, including transaction amount, transaction time, and anonymized features obtained through dimensionality reduction.

Using principal components, some variables were transformed to their numerical variables, V1, V2, … V28 due to security and confidentiality. The only features that were transformed are Time and Amount. The “Amount” feature represents the transaction amount whereas the “Time” feature demonstrates the seconds that elapsed between each transaction. The response variable, “Class,” is categorised to a numerical value of 1 in case of fraud and 0, in genuine transactions. Datasets such as these provide an excellent resource for training and evaluating models.

## Data Inspection and Cleaning

### Basic Exploratory Data Analysis

In this analysis, the dataset was explored to know more about it by familiarising ourselves with the dataset to be working with. It is an essential step because it enables us to obtain a snippet of the features including the number of observations, the type of variables and how the data points were distributed. Understanding these fundamentals will not only set us up for a successful project but also assist us have deep knowledge of the dataset.

Exploring the data helped us identify if there were any missing data. Investigating the dataset allows us to discover missing values, outliers and other data anomalies that might affect the analysis. By detecting these issues early, data scientists can take corrective actions and ensure the accuracy and reliability of the analysis. Basic data exploration analysis assists us to ascertain the relationships and patterns. Exploring the dataset can help us identify relations and patterns that we may not be aware of. This is relevant in hypothesis testing. To clean and process the data, we need to explore it before. For instance, for missing data, we can decide to impute the missing values or remove them from the dataset through exploration analysis. Finally, we can improve our model accuracy by validating assumptions we make about the data such as whether the classes are normally distributed or there is a significant difference between the classes and whether there are any confounding variables that are required to be accounted for.

### Data Preprocessing

Preprocessing data is an integral part of the process of preparing a dataset for data analysis. In preparation for analysis and modelling, data preprocessing is crucial. To ensure data quality and the algorithms are compatible with it a number of preprocessing techniques are used. A data inspection routine was done to deal with missing values, outliers, and inconsistencies as a first step. Whenever possible, missing values are imputed using appropriate techniques, such as mean imputation, or replaced with indicator variables. However, missing values were not present in our dataset.

Figure 6‑2*. Visualising missing values*

A picture containing screenshot, text, line, rectangle

Description automatically generated

We can observe from Figure 6-2 that all the observations for the variables were present. The dataset was examined for outliers and inconsistencies once the data inspection routine was completed.

### Statistical Analysis

Statistical analysis techniques, including box plots and histograms, were used to identify any observations that deviated significantly from the majority of the data. Outliers were treated depending on the nature of the data and the specific models being employed (Smiti, 2020). We can perceive the distribution of the amount in our dataset to be highly skewed. We therefore performed a statistical transformation by applying power transform to render the data points evenly distributed as shown on the right in Figure 6-3. Furthermore, we can infer from the distribution curve that the transactions made were lower, hence, the lower fraud transactions in the dataset. Therefore, there is a possibility of fraud if transactions made with credit cards are higher or huge.

Moreover, further investigation revealed that most of the variables in the dataset were highly skewed, hence, the presence of outliers. Figure 6-4 justifies this analysis. The distributions of data points in variables (transformed into their PCA form) such as V10, V12 and V4 are skewed to the left for the first two and highly to the right respectively whereas V11, V14 and V17 are evenly distributed.

*Figure 6‑3. Distribution of Amount*

A picture containing screenshot, text, plot, diagram

Description automatically generated

*Figure 6‑4. Outliers Detection with Histogram*

A picture containing text, screenshot, diagram, plot

Description automatically generated

To investigate further to validate the presence of outliers, we used a boxplot to graphically identify outliers in the dataset. Boxplot in this instance provided a visual representation of the distribution of the and helped in detecting values that are significantly different from the majority of the observations Walker et al., (2018). The boxplots show various relevant characteristics relating to the outliers such as the skewness and symmetry. We can get an idea of how skewed the data points in each variable are by observing where the median is l situated within each of the boxes. That is, on the one hand, we get the impression that the data is left-skewed if the median is located nearer the lower quartile. On the other hand, the distribution of the data points is skewed to the right if the median is closer to the upper quartile.

Utilising the symmetry of the boxplot, we can prove the overall distribution of the data points based on the data. Outside the whiskers of the boxplots lies the outliers in the dataset as observed in Figure 6-5.

*Figure 6‑5. Detecting Outliers with Boxplot*

A picture containing screenshot, line, font, diagram

Description automatically generated

From Figure 6-5, the following analysis is observed:

Negative skewness (-0.1/lower): Negative skewness suggests that the whiskers of the distributions extend to the left side. This means that the data is left-skewed or negatively skewed, as represented by the median being closer to the upper quartile (the right side of the box) and the whisker on the left side of the box (lower values) being longer than the whisker on the right side (higher values) as we can see from V12 and V17. The data has a concentration of values towards the higher end and some lower extreme values.

Neutral Skewness (Values between 0 and 1): A skewness value within this range suggests that the distribution is relatively symmetrical. The implication is that the data is neither strongly right-skewed nor left-skewed. Within the boxplots of V13 and V15, the median is roughly centred in the middle of the box, indicating a balanced distribution of values on both sides. The length of the whiskers on both sides are almost similar indicating an evenly spread of data points in both directions.

Positive skewness: This skewness suggests that the whiskers of the distributions extend to the right side. This means that the data is right-skewed or positively skewed, as represented by the median being closer to the lower quartile (the right side of the box) and the whisker on the right side of the box (higher values) being longer than the whisker on the left side (lower values) as we can see from V10. Here, the data points have a concentration of values towards the lower end and some higher extreme values.

### Dealing with Outliers with Interquartile Range (IQR)

The interquartile range (IQR) is a statistical instrument for identifying and handling outliers. It measures the spread or dispersion of a dataset by calculating the variance between the upper quartile (75th percentile) and the lower quartile (25th percentile) (Nadim *et al.*, 2019). The extreme values affect the interquartile range lesser than other measures such as range and standard deviation. by capturing the middle 50% of the dataset. The Interquartile Range is applied to deal with outliers using fences.

In this approach, any data point that falls beneath the down fence (lower quartile - 1.5 IQR) *or* exceedsthetopfence(upper quartile, 1.5 IQR) is considered an outlier and can be treated accordingly. These fences provide thresholds beyond which data points are considered potential outliers. When detecting fraud, outliers can be dealt with in many ways. One common approach is to flag and investigate transactions that fall outside the fences as potential fraudulent cases. These transactions can then undergo further scrutiny, such as manual review or additional verification steps, to determine their legitimacy. Additionally, outliers can be used to improve the performance of fraud detection models by incorporating them as part of the training data and creating robust models such as anomaly detectors. The multiplier of 1.5 in this case is subjective and can be adjusted according to the specific requirements of the fraud detection system. Different multipliers can be used to define wider or narrower boundaries, depending on the desired sensitivity to outliers.

At this point, it is evident that our analysis has revealed the presence of outliers. How should we deal with the outliers? Should we remove the outliers in the dataset? We will have more justification if we know the proportion of the presence of the outliers before we can take this decision. Therefore, a critical investigation to ascertain the total number of outliers is crucial. We can do this by employing Tukey's Interquartile Range (IQR) (Saleem et al. 2021). Anything not in the range of Q1-1.5R and Q3 + 1.5 interquartile range is an outlier and can be removed. His technique is used to detect outliers in univariate distribution for symmetric as well in a slightly skewed dataset.

From Figure 6-5, we could observe that most data points in the variables fall out of the range of Q1-1.5R and Q3 + 1.5 interquartile range establishing the fact that there are outliers in our dataset. Applying Tukey’s interquartile range test, we found **31904 outliers** from our analysis that were detected and subsequently deleted. Figure 6-6 gives a visual structure of our dataset after removing outliers from our dataset.

*Figure 6‑6. Dataset structure after removing Outliers*

A picture containing text, screenshot, number

Description automatically generated

Notwithstanding, the higher presence of outliers we will not delete or remove them. This is because every dataset is different (Baesens et al., 2021) and as much as fraud detection is concerned, outliers can help us solve some business questions like “Does transacting a huge amount being an anomaly be a red flag for fraud?” Moreover, as our dataset is not balanced, removing these outliers might compromise the performance of some models. Also, the problem of imbalance in our dataset still exists as demonstrated in Figure 6-6.

There are 203351 minority classes as against 26 of the minority class after removing outliers. Overall, the methodology employed in this dissertation encompasses data preprocessing techniques to ensure data quality and compatibility with machine learning algorithms, strategies to handle data imbalance, and performance evaluation metrics tailored to the specific problem of fraud detection.

Usually, after finding outliers, they are either treated or removed according to their impact on the analysis. In this case, we did nothing to the outliers because they can help us detect anomalies.

Feature scaling is another important step in the preprocessing stage. Considering that the dataset contains features of different scales, they must be scaled to the same range. A number of scaling techniques are commonly employed, including standardization (mean centering and scaling to unit variance) and normalization (scaling to a specific range of values).

In addition, categorical features are encoded into numerical representations to facilitate the implementation of machine learning algorithms based on them. This is accomplished using encoding techniques such as one-hot encoding and label encoding. In our instances, the categorical features were already encoded being our target variable.

## Data Imbalanced and Strategies for Handling it

In a nutshell, we have demonstrated in this study that when there are significant proportions of observations belonging to one class than the other, it is considered to be imbalanced data. This is typically the case in classification machine learning. That is, the class distribution is severely skewed. As an example, the minority class might have 2:200 or 20:2000 observations compared with the majority class. Machine learning models are heavily impacted by this imbalance because they tend to favour the majority class, the genuine transactions while failing to accurately detect the minority class (fraudulent transactions). This is the case of our dataset, in which out of 284,807 transactions, only 492 represent fraudulent transactions whereas 284,315 were genuine transactions.

*Figure 6‑7. Dataset Class Count*

A picture containing text, screenshot, font, design

Description automatically generated

As observed from Figure 6-7, our dataset is highly skewed. The majority class (Genuine Transactions) represents 99.83% whilst the minority class (Fraudulent Transactions) is made up of 0.17% as shown in Figure 6-8 below.

*Figure 6‑8. Proportion of Class Distribution*

A picture containing text, screenshot, rectangle, design

Description automatically generated

*Table 1. Distribution of Classes in the dataset*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Key | Count | Percentage |
| No Fraud | **0** | **284,315** | **99.83%** |
| Fraud | **1** | **492** | **0.17%** |

Generally, there are several strategies for solving class imbalance issues. In this project, the following strategies were employed: random oversampling and undersampling (To be discussed in detail in the next chapter).

## Machine Learning Modelling

This section describes the algorithms deployed in the project. The section also justifies why were selected those models to work with.

### Overview of the Models Used

A range of models we employed to detect fraud in this study. Models selected for this study included a variety of algorithms known for their efficacy in handling fraud detection tasks. These classifiers were carefully selected based on their usability for the task and their potential to provide accurate and reliable predictions. The Random Forest algorithm (Xuan et al., 2018b) was used because it is known to handle large datasets and effectively capture complex relationships in data. Extreme Gradient Boost (XGBoost) is another powerful gradient boosting method (Nijwala *et al.*, 2023)used, which excels at handling imbalanced data and generating highly accurate predictions. As part of the set of models, logistic regression, a classic and interpretable algorithm (Itoo et al., 2021), was also included, offering insights into the impact of different features on fraud likelihood.

Additionally, the Decision Tree algorithm (Patil et al. 2018) was utilised, known for its effortlessness and ability to uncover important decision rules. Also, its ability to estimate the feature which contributed most to prediction was another key factor. The K-nearest neighbour (KNN) model (Raghavan et al., 2019) relied on the principle of similarity to identify fraudulent patterns. The stochastic gradient descent algorithm (Mrozek, 2020), a variant of gradient descent optimisation, has successively been applied to increase the performance of the classifier. A thorough evaluation of the effectiveness of each of these models in detecting credit card fraud is possible through an assessment of their unique advantages and trade-offs. By utilising a diversity of models, the study examines different modelling approaches and their impact on fraud detection accuracy. The selection of these classifiers was driven by their established performance in fraud detection tasks, ensuring a vigorous and comprehensive analysis of payment and fraud risk detection.

### Selection of Model and Justification

The process of selecting the appropriate models for detecting fraud risk involved careful consideration of several factors. The ability of the model to handle the characteristics of the dataset, their interpretability, and their performance in fraud detection tasks were the selection criteria. Random Forest, XGBoost, Logistic Regression, Decision Tree, K-Nearest Neighbours (KNN), and Stochastic Gradient models were chosen. We selected Random Forest and XGBoost since they are capable to accommodate large datasets, capture complex relationships, and handle imbalanced data effectively (Venkatanagendra et al. 2019; Le et al. 2020). Logistic regression was included for its interpretability, allowing us to understand the impact of different features on fraud likelihood (Stiglic *et al.*, 2020). For its simplicity and competence to make assumptions depending on the categorical classes (Charbuty et al., 2021) we chose Decision Tree.

Lastly, the Stochastic Gradient Descent classifier, with its gradient descent optimization, was employed to optimize performance (Bottou et al., 2018). The selected models collectively provide a diverse set of approaches to tackle the credit card fraud problem. They consider factors such as accuracy, interpretability, scalability, and handling of imbalanced data. The study aims to assess the most efficient technique for identifying fraudulent transactions by employing this ensemble of classifiers.

### Splitting the Data and Scaling

To begin with our modelling, we split the dataset into train and test data. However, we will use stratified splitting to avoid data leakage. 80% train - 80% of the data will be used to train the model. 20% test - the remaining 20% will be used to validate our model. We want to keep the same proportions of classes in each split. Using random splitting might render the proportions of the target variable among the different splits to differ from each other. For this reason, we chose to use stratified splitting (Jovanovic *et al.*, 2022) in this study. Additionally, was applied feature scaling ensuring that all variables were on a similar scale. This is relevant because many models are sensitive to the weight of the features. (“Data Normalisation in Data Mining”) Common scaling techniques include standardisation (mean normalization) and normalization (min-max scaling). The choice of scaling technique highly depends on the specific requirements of the models being employed. In this study, we employed the RobustScaler (Raju *et al.*, 2020).

### Resampling techniques

Oversampling techniques (Bae *et al.*, 2021), such as random oversampling and synthetic minority oversampling technique (SMOTE), have been employed by many studies (Jadhav *et al.*, 2022) to increase the representation of the minority class (fraudulent transactions). Undersampling techniques, such as Random Undersampling and NearMiss, have also been utilised in numerous machine learning studies to reduce the number of majority class (non-fraudulent transactions) samples. A combination of oversampling and undersampling techniques was also explored to achieve a balanced dataset in this study.

#### Oversampling Technique

Notwithstanding the effectiveness of the synthetic minority oversampling technique, this study proposes utilising Adaptive Synthetic (ADASYN) algorithm technique. Unlike the former, the adaptive synthetic technique robustly generates different numbers of samples based on estimating the local distribution of the minority class. It further prevents data points from linearly correlating to each other, hence, solving data leakages and enhancing the performance of classifiers. Additionally, it creates a normal distribution between each data point and resamples the minority class adding variance to stop the classes from being characterised by linear relationships (Xu et al., 2020).

#### Undersampling Technique

For undersampling, we propose two techniques: Random Sampling by Dimensionality Reduction; and NearMiss techniques. Dimensionality reduction (Anowar et al., 2021) refers to decreasing the number of random variables in an imbalanced dataset by obtaining a set of principal components of variables.

Studies have proposed using a lot of methods in which complex problems are simplified, eliminating redundancy and also reducing the potential of model overfitting. In this study, we used Rus Scalar to reduce the classes making each class even (50% each).

A picture containing text, font, screenshot, line

Description automatically generated

*Table 2. Dimension of Dataset after Applying Rus*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Key | Count | Percentage |
| No Fraud | **0** | **492** | **50%** |
| Fraud | **1** | **492** | **50%** |

We then applied a dimensionality reduction technique implementing T-distributed Stochastic Neighbourhood Embedding (t-SNE), Principal Component Analysis (PCA) and Truncated Singular Values Decomposition (TruncatedSVD). That is, these techniques were used to remove data points to have a balanced dataset and to avoid overfitting.

To determine which feature influences most whether a specific case will be an instance of default, we used a correlation matrix to view relationships between data points after the reduction. In ascertaining which feature has the highest correlation, it is relevant to the right sample. From Figure 6-9 we can observe the correlation of a dataset with a higher dimensionality that contains an enormous number of columns or variables with large observations.

*Figure 6‑9. Correlation Matrix Before Reduction*

A picture containing screenshot, colorfulness, text, rectangle

Description automatically generated

*Figure 6‑10. Undersampling by Dimensionality Reduction*

A picture containing text, screenshot, diagram, map

Description automatically generated

From Figure 6-10, we can observe that, the data points have been reduced by applying T-distributed Stochastic Neighbourhood Embedding (t-SNE), Principal Component Analysis (PCA) and Truncated Singular Values Decomposition (TruncatedSVD) respectively. These reduction strategies have successfully reduced the large dimension of the data points avoiding the risk of learning wrong data patterns and the features selected should remove insignificant attributes and aspects of fraud class (Anowar et al., 2021).

The application of t-distributed stochastic neighbourhood embedding preserved the location structure of the majority class whilst reducing it, the principal component analysis resolves the feature selection issues from the point of numerical analysis by locating the most appropriate number of principal components (Chari et al., 2022). The use of TruncatedSVD produces a factorisation where the number of columns is the same as the specified truncated. By observing Figure 6-11, we can see that the correlation matrix has also been reduced in dimension with most of the data relationships removed.

*Figure 6‑11. Correlation after Dimension Reduction*

A picture containing screenshot, text, plot, diagram

Description automatically generated

#### Undersampling by NearMiss

Proposed by Jianping Zhang (2003) as cited by Xiao et al. (2019), NearMiss refers to a system of undersampling methods that selects instances depending on the majority class’s distance instances to the minority class instances. Three versions of this strategy namely NearMiss version-1, NearMiss version-2 and NearMiss version-3 are available. In this study, we used version-3.

NearMiss Version-3 chooses a given number of the majority class instances for each instance in the closest minority class. Given that it will only retain those majority class instances that are on the decision boundary, this version appears desirable.

*Table 3. Dimension of Dataset after Applying NearMiss Version-3*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Key | Count | Percentage |
| No Fraud | **0** | **359** | **48%** |
| Fraud | **1** | **378** | **52%** |

It is worth noting that, whilst resampling potentially prevents data leakages, underfitting and overfitting, it also has adverse effects on the performance of the models. Whereas oversampling data had better model performance, undersampling data did worse. This is further illustrated in the next Chapter.

### Hyperparameter Tuning Techniques

Hyperparameter tuning plays a significant role in the optimisation of models' performance. It is the process of tuning the parameters present as the tuples when building models (Agrawal, 2021). Machine learning classifiers do not learn these parameters; hence, we have to define them. It is essential to fine-tune the hyperparameters of the chosen models to achieve optimal results in the context of credit card fraud detection. Hyperparameter tuning can be accomplished by using different techniques and some are outlined in this section.

Grid search and random search were the primary hyperparameter tuning techniques. Grid search involves systematically searching through a predefined grid of hyperparameter combinations and evaluating each combination's performance (Liashchynskyi et al., 2019). Based on predetermined evaluation metrics, this technique allows us to explore various hyperparameter values and find the most appropriate combination. Except for Stochastic Gradient Descent which we used hyperopt for optimising, all the models used gridsearch. Hyperopt optimisation can perform model hierarchical hyper-parameters for various classifiers (Yang et al., 2020). Random search, on the other hand, randomly samples hyperparameter values from predefined distributions and evaluates their performance (Probst et al., 2019). When the hyperparameter space is large, this technique provides a more efficient search strategy than grid search.

To guarantee a comprehensive search for optimal hyperparameters, the process of tuning integrates various evaluation metrics. For instance, the area under the receiver operating characteristic curve (AUC-ROC) was employed as a performance metric to assess the models' ability to differentiate between non-fraudulent and transactions as approached by Fawcett (2006) cited in (Muschelli, 2020). Additionally, while minimizing false positives (Moodley *et al.*, 2020), precision, recall, and F1-score were utilised to assess the performance of the models in correctly identifying fraudulent instances.

### Cross Validation

Cross-validation was used to estimate the generalisation performance of the model and mitigate overfitting (Itoo et al., 2021) as part of the hyperparameter tuning procedure. With multiple folds of the dataset serving as both training and validation data, cross-validation facilitates the evaluation of different hyperparameter settings with more reliability. The best way to cross validate our data is during cross-validation. This is done so that our training data will not be contaminated or compromised with the sampled data.

The best model is the one that makes better predictions on the unseen or new data and avoids overfitting and/or underfitting but not the one that gives accurate predictions on the training set. Hence, the need for cross-validation. Cross-validation can be referred to as a resampling method employed to evaluate a model on limited sampled data (Santos *et al.*, 2018). There is only one parameter identified as K in this method, which indicates the number of groups in which the sample data is divided. This method can also be regarded as K-fold cross-validation. Its main function is to test whether a machine learning classifier is capable of making a good prediction on new data or data it has not seen before.

To identify the optimal combination of hyperparameters for each selected model, we combined grid search, random search, hyperopt and evaluation metrics as well as cross-validation techniques. In an attempt to enhance the ability to detect credit card fraud and ensure its effectiveness in real-world situations, a continuous iterative process is used.

### Performance Evaluation Metrics

While traditional evaluation metrics like accuracy and confusion matrix can be useful for assessing a model's performance, they can be inaccurate when applied to imbalanced datasets (Najadat *et al.*, 2020). Performance metrics that are more suitable for imbalanced datasets are discussed in this section.

*Table 4. Traditional Confusion Matrix*

|  |  |  |
| --- | --- | --- |
| **Predicted** | **Actual** | |
| **No Fraud** | **Fraud** |
| No Fraud | True Negative (TN) | False Negative (FN) |
| Fraud | False Positive (FP) | True Positive (TP) |

#### Precision and Recall

Precision and recall metrics are commonly used in imbalanced classification problems. The precision of a prediction represents the ratio of correctly predicted positive instances out of all predicted positive instances, while the recall measures the proportions of correctly predicted positive instances out of all actual positive instances. Whilst a high precision signifies a low false positive rate, a high recall indicates a low false negative rate (Rtayli et al., 2020). These metrics are particularly useful in fraud detection scenarios, as we want to minimize both false positives (genuine transactions classified as fraudulent) and false negatives (fraudulent transactions classified as legitimate).

Precision or Positive Prediction Value: The measure of True Positive results in all positive predictions. Can be expressed as:

Precision = TP/(TP + FP): produces the accuracy on cases predicted as fraud

Recall: Measures all correct positive predictions made out of all positive predictions.

Recall = TP/(TP +FN): produces the accuracy on actual cases predicted as fraud

#### F1 Score

The F1 score is a harmonic mean of precision and recall. The metric balances precision and recall. F1 scores can be useful when we are seeking a balance between identifying fraudulent transactions accurately and minimizing false alarms. Considering both false positives and false negatives, it is a suitable metric for datasets that are imbalanced (Yee et al., 2018).

F1 score = 2 x Precision Recall/(Precision + Recall)

#### Precision-Recall Curve (AUPRC)

Precision-Recall Curve (AUPRC): Depending on the classification threshold, the precision-recall curve represents the trade-off between precision and recall. The area under this curve summarizes the overall performance of the model. For imbalanced datasets, and in particular, when the class distribution is highly skewed, it provides a robust evaluation metric.

#### Receiver Operating Characteristic Curve – Area Under Curve (ROC-AUC)

In addition to the Receiver Operating Characteristic (ROC) curve, the area under the curve (AUC-ROC) plots the true positive rate (TPR) against the false positive rate (FPR). Specifically, the Receiver Operating Characteristics (ROC) is a probability whereas Area Under Curve is a representation of the measure of separability. It Illustrates how much a model can distinguish between minority and majority classes. The AUC-ROC gives a comprehensive evaluation of the model's performance at different threshold settings by measuring the area under this curve. Although the AUC-ROC is useful in imbalanced datasets, it can be misleading in minority-class datasets because it is less sensitive to minority performance.

A higher area under the curve (AUC) score signifies how accurate a model is at predicting 1 (Fraud) as 1 and 0 (non-Fraud) as 0. A well-performing model has an AUC closer to 1 (Zhang *et al.*, 2021), which means it has a better measure of separability. If a model performed poorly then it has an AUC near 0 which means it has the worst evaluation of separability. When AUC is 0.8, it means there is an 80% chance that the model will be able to distinguish between and negative class positive class. A 0.5 AUC score means that the model has no distinction ability to differentiate between non-fraud and fraud classes. The model is predicting a negative class as a positive class and vice versa when AUC is approximately 0.

Besides these metrics, it is important to consider domain-specific factors and business requirements when evaluating fraud detection models. Developing customized evaluation metrics can take specific needs into account, such as optimizing for a particular cost function or accounting for false positives and false negatives. Overall, the model's performance in detecting credit card fraud can be evaluated using precision, recall, F1 score, AUPRC, ROC-AUC and customized evaluation metrics, considering the challenges posed by imbalanced datasets.

*Table 5. Summary of Performance Metrics*

|  |  |
| --- | --- |
| **Metric** | **Formula** |
| Precision | TP/(TP + FP) |
| Recall | TP/(TP + FN) |
| F1 Score | 2 x (Precision x Recall)/(Precision + Recall) |
| Negative Positive Rate | TN/(TN + FN) |
| False Positive Rate | FP/(FP + FN) |
| True Positive Rate | TP/(TP + FN) |
| ROC | TP Rate against FP Rate |
| Accuracy  (Not Used Here) | TN + TP/TP + FP + FN + TN |

# Chapter 4: Results and Discussions

This chapter is a presentation of our experiments results comparing the performance evaluation of Logistic Regression (LR), K-Nearest Neighbour (KNN), Decision (Extra) Tree, Extra Gradient Boost (XGboost), Random Forest (RF), XGBRandom Forest (XGBRF), Extra Trees and Stochastic Gradient Descent (SGD). As already emphasised, because of the challenge of an imbalanced dataset, accuracy and the confusion matrix are not used as a metric, however, we masked AUC score on each of the matrix just for observation and comparison purposes. Instead, precision, recall, F1 score, Precision-Recall (AUPRC) and ROC-AUC are the focus here especially Area Under Curve (AUC).

To start with, we built the main model from the original dataset with the default parameters and thereafter we combined all the classifiers to vote on predictions for us. Thus, a voting classifier is a technique that put multiple classifiers to make a final prediction. This is also known as an ensemble approach that can enhance the performance of models and also ensures the stability of predictions, most especially when models are susceptible to error and/or biases. Another ensemble approach used to optimise the main model is the stacking algorithm. It combines multiple classifiers to create one powerful model where the combined classifiers learn from the strength of each classifier.

For much better performance, we optimised each model with GridSearch except for Stochastic Gradient Descent, for which we used Hyperopt.

## Main Model with Default Parameters

*Table 6. Results from Main Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL NAME** | **PRECISION** | **RECALL** | **PRECISION-RECALL** | **AUC** |
| Random Forest | 0.96 | 0.71 | 0.80 | 0.92 |
| Decision Tree | 0.99 | 0.73 | 0.82 | 0.94 |
| XGBoost | 0.97 | 0.75 | 0.80 | 0.96 |
| XGBRandom Forest | 0.95 | 0.74 | 0.80 | 0.97 |
| KNN | 0.96 | 0.67 | 0.73 | 0.51 |
| Logistic Regression | 0.86 | 0.59 | 0.70 | 0.46 |
| SGD | 0.80 | 0.33 | 0.64 | 0.50 |

*Figure 7‑1. AUC Score Plot from Main Model*

A picture containing text, screenshot, line, diagram

Description automatically generated

*Figure 7‑2. Stack Classifier for Mian Model*

A picture containing text, screenshot, plot, line

Description automatically generated

Analysing Table 6, figure 8-1 and 8-2, we could observe that whilst some models performed extremely well others performed poorly. XGBoost had the highest recall scores, whereas Decision Tree performed well when precision-recall is used as a metric and XGBoost Random Forest Classifier performed extremely well when AUC is used as a metric. When all the models are combined, they performed tremendously with AUC of 0.97 and predicted each class correctly.

## Optimised Main Model

*Table 7. Results from Optimised Main Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL NAME** | **PRECISION** | **RECALL** | **PRECISION-RECALL** | **AUC** |
| Decision Tree | 0.80 | 0.74 | 0.65 | 0.91 |
| XGBoost | 0.97 | 0.72 | 0.83 | 0.98 |
| KNN | 0.97 | 0.68 | 0.77 | 0.88 |
| Logistic Regression | 0.34 | 0.81 | 0.70 | 0.50 |
| SGD | 0.76 | 0.77 | 0.64 | 0.50 |

*Figure 7‑3 AUC Score Plot for Optimised Main Model*

A picture containing text, screenshot, line, diagram

Description automatically generated

*Figure 7‑4. Stack Classifier for Optimised Mian Model*

A picture containing text, screenshot, plot, line

Description automatically generated

We can observe from Table 7 that whilst XGBoost and KNN had the highest precision scores, Logistic Regression performed well when recall is used as a metric and XGBoost again performed extremely well when AUC and Precision-Recall are used to measure performance. When we combine all the models, they performed tremendously with AUC of 0.97 and predicted each class correctly using the stacking algorithm.

## Oversample Model

*Table 8. Results from Oversample Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL NAME** | **PRECISION** | **RECALL** | **PRECISION-RECALL** | **AUC** |
| Random Forest | 0.92 | 0.76 | 0.80 | 0.94 |
| Decision Tree | 0.94 | 0.76 | 0.83 | 0.95 |
| XGBoost | 0.77 | 0.80 | 0.80 | 0.96 |
| XGBRandom Forest | 0.03 | 0.87 | 0.66 | 0.95 |
| KNN | 0.36 | 0.82 | 0.50 | 0.50 |
| Logistic Regression | 0.02 | 0.89 | 0.61 | 0.50 |
| SGD | 0.01 | 0.90 | 0.69 | 0.50 |

*Figure 7‑5. AUC Plot for Oversample Model.*

A picture containing text, screenshot, line, diagram

Description automatically generated

*Figure 7‑6. AUC Plot for Oversample Voting Model*

A picture containing text, screenshot, plot, line

Description automatically generated

*Figure 7‑7. AUC Plot for Oversample Stacking Model*

A picture containing text, diagram, plot, line

Description automatically generated

The models on the oversample data generally perform well. However, KNN, Logistic Regression and SGD performed abysmally. This can be a result of their inability to handle a large amount of imbalanced data. The voting ensemble performed extremely better.

## Optimised Oversample Model

*Table 9. Results from Optimised Oversample Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL NAME** | **PRECISION** | **RECALL** | **PRECISION-RECALL** | **AUC** |
| Decision Tree | 0.01 | 0.83 | 0.2 | 0.92 |
| XGBoost | 0.82 | 0.79 | 0.82 | 0.97 |
| KNN | 0.5 | 0.79 | 0.52 | 0.90 |
| Logistic Regression | 0.01 | 0.91 | 0.66 | 0.50 |
| SGD | 0.01 | 0.96 | 0.13 | 0.50 |

*Figure 7‑8.AUC Plot for Optimised Oversample Model*

A picture containing text, screenshot, diagram, line

Description automatically generated

*Figure 7‑9. AUC Plot for Optimised Oversample Voting Classifier*

A picture containing text, screenshot, line, diagram

Description automatically generated

The voting Classifier gave a tremendous performance but was not better than the actual oversample data. KNN with the optimised oversample data performed far better than the actual oversample data.

## Undersample Model

*Table 10. Results from Undersample Model*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **MODEL NAME** | **PRECISION** | **RECALL** | **PRECISION-RECALL** | **AUC** |
| Random Forest | 0.63 | 0.75 | 0.75 | 0.92 |
| Decision Tree | 0.94 | 0.76 | 0.83 | 0.77 |
| XGBoost | 0.03 | 0.78 | 0.70 | 0.92 |
| XGBRandom Forest | 0.41 | 0.75 | 0.73 | 0.90 |
| KNN | 0.36 | 0.82 | 0.50 | 0.67 |
| Logistic Regression | 0.02 | 0.89 | 0.61 | 0.50 |
| SGD | 0.01 | 0.90 | 0.69 | 0.50 |

*Figure 7‑10. AUC Plot of Undersample Data*

A picture containing text, screenshot, diagram, line

Description automatically generated

The model could capture every detail because there was not much weight and subsequently. Generally, models performed poorly on the undersampled data. Notwithstanding, the stacking classifier did well with the undersampled data.

*Figure 7‑11. AUC Plot of Undersample Stacking Classifier*

A picture containing text, screenshot, line, plot

Description automatically generated

## Interpreting Model Results (Explainable AI)

To understand these complex black-box models, we rely on the Local Interpretable Model-Agnostic Explanations (LIME) to generate local explanations for individual predictions. It is done by generating explanations for why a particular transaction was flagged as fraudulent or not. Local Interpretable Model-Agnostic Explanations flags indicate and flags the most important features that influence fraud. From Figures 7-12 and 7-13, the algorithm flags V2, V4 and Amount as the most important features (in orange) influencing the occurrence of fraud. These features were above a certain threshold, hence, they were flagged as fraud. The model has detected an anomaly or unusual behaviour. For example, Amount was flagged as a fraudulent feature because the model had 75% of unusual behaviour.

*Figure 7‑12. Explainable AI - Features that Model Flags as Fraud*

A picture containing text, number, font, diagram

Description automatically generated

A picture containing text, screenshot, number, software

Description automatically generated

*Figure 7‑13. Explainable AI - Features that Model Flags as No Fraud*

A picture containing text, number, software, diagram

Description automatically generated

Although this is the most effective approach to explaining model results, another way to add interpretability to models is to analyse feature importance within some algorithms. For instance, in this study, the permutation feature importance from the Random Forest Classifier by shuffling randomly the values of each feature and also measuring the outcome as they decrease in the performance of the model can enable us to understand our results using feature importance. Figure 7-14, when carefully observed indicates that V10 is the most important feature and V23 is the least influencing fraud transactions or not.

*Figure 7‑14. Random Forest Feature Importance*

A picture containing text, screenshot, plot, diagram

Description automatically generated

## Conclusion and Discission

This research examined the application of machine learning techniques for fraud detection in financial transactions. We explored the appropriate methods to handle an imbalanced dataset and industry-standard data reprocessing techniques. Our analysis revealed that optimised models, especially those trained on oversampled datasets achieved the best performance compared to the models built with undersampled data. However, we must guard against data leakages. In conclusion, our research demonstrates the potential of machine learning techniques in detecting fraud. Nonetheless, there are opportunities for further study to address the challenges of imbalanced datasets although this study has demonstrated efficient ways of ensuring data balance. We focused on increasing the minority class (fraud instances) to have the same or a closer weight as the majority class. In our study, we proposed using Adaptive Synthetic (ADASYN) algorithm technique because it is capable of preventing data leakage by ensuring the data points are not linearly correlated. We also decreased the majority class (non-fraud instances) to balance the data by applying dimensionality reduction and NearMiss. Both techniques were effective, however, classifiers performed poorly on the undersampled dataset. Future studies need to research the best practices for balancing data.

Again, future studies should explore new evaluation metrics and model interpretability. Further research could investigate the use of sophisticated ensemble methods, and anomaly detectors expanding more transparent ways of incorporating explainable artificial intelligence to understand model decisions and results. These future studies should also focus on real-time fraud detection systems and explore applying this to a larger scale of financial datasets.

### Critical Analysis of This Study

This research encompasses all the data science skillsets and techniques to analyse data, identify trends and patterns to provide insights and leverage predictive models to support data-driven decision-making. Nevertheless, this study could have been more enlightened if variables in the dataset were complete and not transformed into their numerical forms. This affected the analysis, especially during the advanced exploratory data analysis as well as interpreting how the models made their predictions. Again, the research focuses solely on implementing machine learning techniques and algorithms for fraud detection in financial transactions. Meanwhile, it could benefit from considering a broader perspective by incorporating additional data sources or exploring unsupervised techniques including anomaly detection and deep learning.

To critique further, although this research discusses different machine learning algorithms used to detect fraud, it has a narrow comprehensive comparative analysis of their performance. A more in-depth evaluation of the metrics for measuring their performance could have provided a better understanding of their strength and limitations. Moreover, this study primarily focused on the technical aspects of fraud detection using machine learning and therefore lacks real-world application. However, incorporating real-world practical implications and applications of the findings would be of great benefit. This could have been achieved by discussing the integration of the proposed algorithms into existing fraud detection challenges.

To improve this, it is ideal to expand the scope to include unsupervised learning algorithms. Conducting a comparative analysis, providing more detailed explanations of the model results and feature selections as well as considering a real-world application should be the focus of future studies. By focusing on these, the research will be enhanced and its practical relevance can be further emphasised.

# Reference List

Agrawal, T. (2021) ‘Hyperparameter Optimization in Machine Learning’, *Hyperparameter Optimization in Machine Learning* [Preprint]. Available at: https://doi.org/10.1007/978-1-4842-6579-6.

Alawida, M. *et al.* (2022) ‘A deeper look into cybersecurity issues in the wake of Covid-19: A survey’, *Journal of King Saud University - Computer and Information Sciences*, 34(10), pp. 8176–8206. Available at: https://doi.org/10.1016/J.JKSUCI.2022.08.003.

Anowar, F., Sadaoui, S. and Selim, B. (2021) ‘Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE)’, *Computer Science Review*, 40, p. 100378. Available at: https://doi.org/10.1016/J.COSREV.2021.100378.

Awotunde, J.B., Chakraborty, C. and Adeniyi, A.E. (2021) ‘Intrusion Detection in Industrial Internet of Things Network-Based on Deep Learning Model with Rule-Based Feature Selection’, *Wireless Communications and Mobile Computing*, 2021. Available at: https://doi.org/10.1155/2021/7154587.

Awoyemi, J.O., Adetunmbi, A.O. and Oluwadare, S.A. (2017) ‘Credit card fraud detection using machine learning techniques: A comparative analysis’, in *Proceedings of the IEEE International Conference on Computing, Networking and Informatics, ICCNI 2017*, pp. 2–4. Available at: https://doi.org/10.1109/ICCNI.2017.8123782.

Aziz, S. and Dowling, M. (2019) ‘Machine Learning and AI for Risk Management’, *Palgrave Studies in Digital Business and Enabling Technologies*, pp. 33–50. Available at: https://doi.org/10.1007/978-3-030-02330-0\_3/TABLES/1.

Bae, S.Y. *et al.* (2021) ‘Effective data-balancing methods for class-imbalanced genotoxicity datasets using machine learning algorithms and molecular fingerprints’, *Computational Toxicology*, 20, p. 100178. Available at: https://doi.org/10.1016/J.COMTOX.2021.100178.

Baesens, B., Höppner, S. and Verdonck, T. (2021) ‘Data engineering for fraud detection’, *Decision Support Systems*, 150, p. 113492. Available at: https://doi.org/10.1016/J.DSS.2021.113492.

Bhattacharyya, S. *et al.* (2011) ‘Data mining for credit card fraud: A comparative study’, *Decision Support Systems*, 50(3), pp. 602–613. Available at: https://doi.org/10.1016/J.DSS.2010.08.008.

BoG (2022) *BANKING, SDI AND EMI FRAUD REPORT FINANCIAL STABILITY DEPARTMENT 2*. Accra. Available at: https://www.bog.gov.gh/wp-content/uploads/2022/06/2021-FRAUD-REPORT-Industry-Version-For-Publication.pdf (Accessed: 5 May 2023).

Bottou, L., Curtis, F.E. and Nocedal, J. (2018) ‘Optimization Methods for Large-Scale Machine Learning’, *https://doi.org/10.1137/16M1080173*, 60(2), pp. 223–311. Available at: https://doi.org/10.1137/16M1080173.

Burnes, D., DeLiema, M. and Langton, L. (2020) ‘Risk and protective factors of identity theft victimization in the United States’, *Preventive Medicine Reports*, 17, p. 101058. Available at: https://doi.org/10.1016/J.PMEDR.2020.101058.

Carcillo, F. *et al.* (2021a) ‘Combining unsupervised and supervised learning in credit card fraud detection’, *Information Sciences*, 557, pp. 317–331. Available at: https://doi.org/10.1016/J.INS.2019.05.042.

Carcillo, F. *et al.* (2021b) ‘Combining unsupervised and supervised learning in credit card fraud detection’, *Information Sciences*, 557, pp. 317–331. Available at: https://doi.org/10.1016/J.INS.2019.05.042.

Caroline Cynthia, P. and Thomas George, S. (2021) ‘An outlier detection approach on credit card fraud detection using machine learning: A comparative analysis on supervised and unsupervised learning’, in *Advances in Intelligent Systems and Computing*. Available at: https://doi.org/10.1007/978-981-15-5285-4\_12.

Chang, V. *et al.* (2022) ‘Digital payment fraud detection methods in digital ages and Industry 4.0’, *Computers and Electrical Engineering*, 100, p. 107734. Available at: https://doi.org/10.1016/J.COMPELECENG.2022.107734.

Charbuty, B. and Abdulazeez, A. (2021) ‘Classification Based on Decision Tree Algorithm for Machine Learning’, *Journal of Applied Science and Technology Trends*, 2(01), pp. 20–28. Available at: https://doi.org/10.38094/jastt20165.

Chari, T. and Pachter, L. (2022) ‘The Specious Art of Single-Cell Genomics’, *bioRxiv*, p. 2021.08.25.457696. Available at: https://doi.org/10.1101/2021.08.25.457696.

Dal Pozzolo, A. *et al.* (2014) ‘Learned lessons in credit card fraud detection from a practitioner perspective’, *Expert Systems with Applications*, 41(10), pp. 4915–4928. Available at: https://doi.org/10.1016/J.ESWA.2014.02.026.

Dal Pozzolo, A. *et al.* (2015) *Credit Card Fraud Detection | Kaggle*, *Symposium on Computational Intelligence and Data Mining (CIDM), IEEE*. Available at: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud (Accessed: 7 May 2023).

Dornadula, V.N. and Geetha, S. (2019) ‘Credit Card Fraud Detection using Machine Learning Algorithms’, *Procedia Computer Science*, 165, pp. 631–641. Available at: https://doi.org/10.1016/J.PROCS.2020.01.057.

Dümbgen, L. and Riedwyl, H. (2007) ‘On fences and asymmetry in box-and-whiskers plots’, *American Statistician*, 61(4), pp. 356–359. Available at: https://doi.org/10.1198/000313007X247058.

Frigge, M., Hoaglin, D.C. and Iglewicz, B. (1989) ‘Some implementations of the boxplot’, *American Statistician*, 43(1), pp. 50–54. Available at: https://doi.org/10.1080/00031305.1989.10475612.

Hodjat Hamidi, M.S. (2019) ‘A model to analyze the effect of mobile banking adoption on customer interaction and satisfaction: A case study of m-banking in Iran’, *Telematics and Informatics*, 38, pp. 166–181.

IC3 (2021a) ‘2020 IC3 Alabama Annual Report’, *2020 Internet Crime Report* [Preprint]. Available at: https://www.ic3.gov/Media/PDF/AnnualReport/2020State/StateReport.aspx (Accessed: 4 May 2023).

IC3 (2021b) *IC3 Releases 2020 Internet Crime Report — FBI*, *FBI*. Available at: https://www.fbi.gov/news/press-releases/fbi-releases-the-internet-crime-complaint-center-2020-internet-crime-report-including-covid-19-scam-statistics (Accessed: 4 May 2023).

Itoo, F., Meenakshi and Singh, S. (2021) ‘Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection’, *International Journal of Information Technology (Singapore)*, 13(4), pp. 1503–1511. Available at: https://doi.org/10.1007/S41870-020-00430-Y/FIGURES/8.

Jadhav, A. *et al.* (2022) ‘An Empirical Assessment of Performance of Data Balancing Techniques in Classification Task’, *Applied Sciences 2022, Vol. 12, Page 3928*, 12(8), p. 3928. Available at: https://doi.org/10.3390/APP12083928.

Jovanovic, D. *et al.* (2022) ‘Tuning Machine Learning Models Using a Group Search Firefly Algorithm for Credit Card Fraud Detection’, *Mathematics 2022, Vol. 10, Page 2272*, 10(13), p. 2272. Available at: https://doi.org/10.3390/MATH10132272.

Jung, B., Choi, K.-S. and Lee, C. (2022) ‘Dynamics of Dark Web Financial Marketplaces: An Exploratory Study of Underground Fraud and Scam Business’, *International Journal of Cybersecurity Intelligence & Cybercrime*, 5(2). Available at: https://vc.bridgew.edu/ijcic/vol5/iss2/2 (Accessed: 4 May 2023).

Jurgovsky, J. *et al.* (2018) ‘Sequence classification for credit-card fraud detection’, *Expert Systems with Applications*, 100, pp. 234–245. Available at: https://doi.org/10.1016/J.ESWA.2018.01.037.

Khemakhem, S. and Boujelbene, Y. (2018) ‘Support vector machines for credit risk assessment with imbalanced datasets’, *International Journal of Data Mining, Modelling and Management*, 10(2), pp. 171–187. Available at: https://doi.org/10.1504/IJDMMM.2018.092538.

Kim, E. *et al.* (2019) ‘Champion-challenger analysis for credit card fraud detection: Hybrid ensemble and deep learning’, *Expert Systems with Applications*, 128, pp. 214–224. Available at: https://doi.org/10.1016/J.ESWA.2019.03.042.

Kim, J., Kim, H.J. and Kim, H. (2019) ‘Fraud detection for job placement using hierarchical clusters-based deep neural networks’, *Applied Intelligence*, 49(8), pp. 2842–2861. Available at: https://doi.org/10.1007/S10489-019-01419-2/FIGURES/7.

Le, T.T., Fu, W. and Moore, J.H. (2020) ‘Scaling tree-based automated machine learning to biomedical big data with a feature set selector’, *Bioinformatics*, 36(1), pp. 250–256. Available at: https://doi.org/10.1093/BIOINFORMATICS/BTZ470.

Leo, M., Sharma, S. and Maddulety, K. (2019) ‘Machine Learning in Banking Risk Management: A Literature Review’, *Risks 2019, Vol. 7, Page 29*, 7(1), p. 29. Available at: https://doi.org/10.3390/RISKS7010029.

Liashchynskyi, Petro and Liashchynskyi, Pavlo (2019) ‘Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS’. Available at: https://arxiv.org/abs/1912.06059v1 (Accessed: 28 May 2023).

Makki, S. *et al.* (2019) ‘An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection’, *IEEE Access*, 7, pp. 93010–93022. Available at: https://doi.org/10.1109/ACCESS.2019.2927266.

McKinlay, J. (1982) ‘Videotex’, *Australian Library Journal*, 31(3), p. 12. Available at: https://doi.org/10.1080/00049670.1982.10755457.

Mienye, I.D. and Sun, Y. (2023) ‘A Deep Learning Ensemble with Data Resampling for Credit Card Fraud Detection’, *IEEE Access* [Preprint]. Available at: https://doi.org/10.1109/ACCESS.2023.3262020.

Miller, T. (2019) ‘Explanation in artificial intelligence: Insights from the social sciences’, *Artificial Intelligence*, 267, pp. 1–38. Available at: https://doi.org/10.1016/J.ARTINT.2018.07.007.

Mittal, S. and Tyagi, S. (2019) ‘Performance evaluation of machine learning algorithms for credit card fraud detection’, *Proceedings of the 9th International Conference On Cloud Computing, Data Science and Engineering, Confluence 2019*, pp. 320–324. Available at: https://doi.org/10.1109/CONFLUENCE.2019.8776925.

Mohammed, R.A. *et al.* (2018) ‘Scalable machine learning techniques for highly imbalanced credit card fraud detection: A comparative study’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11013 LNAI, pp. 237–246. Available at: https://doi.org/10.1007/978-3-319-97310-4\_27/COVER.

Moodley, R. *et al.* (2020) ‘A product-centric data mining algorithm for targeted promotions’, *Journal of Retailing and Consumer Services*, 54, p. 101940. Available at: https://doi.org/10.1016/J.JRETCONSER.2019.101940.

Mrozek, P., Panneerselvam, J. and Bagdasar, O. (2020) ‘Efficient resampling for fraud detection during anonymised credit card transactions with unbalanced datasets’, *Proceedings - 2020 IEEE/ACM 13th International Conference on Utility and Cloud Computing, UCC 2020*, pp. 426–433. Available at: https://doi.org/10.1109/UCC48980.2020.00067.

Munappy, A.R. *et al.* (2022) ‘Data management for production quality deep learning models: Challenges and solutions’, *Journal of Systems and Software*, 191, p. 111359. Available at: https://doi.org/10.1016/J.JSS.2022.111359.

Muschelli, J. (2020) ‘ROC and AUC with a Binary Predictor: a Potentially Misleading Metric’, *Journal of Classification*, 37(3), pp. 696–708. Available at: https://doi.org/10.1007/S00357-019-09345-1/FIGURES/4.

Nadim, A.H. *et al.* (2019) ‘Analysis of machine learning techniques for credit card fraud detection’, *Proceedings - International Conference on Machine Learning and Data Engineering, iCMLDE 2019*, pp. 42–47. Available at: https://doi.org/10.1109/ICMLDE49015.2019.00019.

Najadat, H. *et al.* (2020) ‘Credit Card Fraud Detection Based on Machine and Deep Learning’, *2020 11th International Conference on Information and Communication Systems, ICICS 2020*, pp. 204–208. Available at: https://doi.org/10.1109/ICICS49469.2020.239524.

Nijwala, D.S. *et al.* (2023) ‘Extreme Gradient Boost Classifier based Credit Card Fraud Detection Model’, *Proceedings - IEEE International Conference on Device Intelligence, Computing and Communication Technologies, DICCT 2023*, pp. 500–504. Available at: https://doi.org/10.1109/DICCT56244.2023.10110188.

Patil, S., Nemade, V. and Soni, P.K. (2018) ‘Predictive Modelling For Credit Card Fraud Detection Using Data Analytics’, *Procedia Computer Science*, 132, pp. 385–395. Available at: https://doi.org/10.1016/J.PROCS.2018.05.199.

Philip Greenspun (2018) *Shopping and banking on a computer network in the 1980s (Minitel) – Philip Greenspun’s Weblog*, *Philip Greenspun’s Weblog*. Available at: https://philip.greenspun.com/blog/2018/03/30/shopping-and-banking-on-a-computer-network-in-the-1980s-minitel/ (Accessed: 3 May 2023).

Probst, P., Wright, M.N. and Boulesteix, A.L. (2019) ‘Hyperparameters and tuning strategies for random forest’, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), p. e1301. Available at: https://doi.org/10.1002/WIDM.1301.

PwC (2022) *PwC’s Global Economic Crime and Fraud Survey 2022*, *Protecting the perimeter: A new frontier of platform fraud*. Available at: https://www.pwc.com/gx/en/services/forensics/economic-crime-survey.html (Accessed: 4 May 2023).

Raghavan, P. and Gayar, N. El (2019) ‘Fraud Detection using Machine Learning and Deep Learning’, *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019*, pp. 334–339. Available at: https://doi.org/10.1109/ICCIKE47802.2019.9004231.

Raju, V.N.G. *et al.* (2020) ‘Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification’, *Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020*, pp. 729–735. Available at: https://doi.org/10.1109/ICSSIT48917.2020.9214160.

Rtayli, N. and Enneya, N. (2020) ‘Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization’, *Journal of Information Security and Applications*, 55, p. 102596. Available at: https://doi.org/10.1016/J.JISA.2020.102596.

Sadikin, F. and Kumar, S. (2020) ‘ZigBee IoT Intrusion Detection System: A Hybrid Approach with Rule-based and Machine Learning Anomaly Detection’, *IoTBDS* [Preprint]. Available at: https://doi.org/10.5220/0009342200570068.

Sahu, A., Gm, H. and Gourisaria, M.K. (2020) ‘A Dual Approach for Credit Card Fraud Detection using Neural Network and Data Mining Techniques’, *2020 IEEE 17th India Council International Conference, INDICON 2020* [Preprint]. Available at: https://doi.org/10.1109/INDICON49873.2020.9342462.

Sailusha, R. *et al.* (2020) ‘Credit Card Fraud Detection Using Machine Learning’, *Proceedings of the International Conference on Intelligent Computing and Control Systems, ICICCS 2020*, pp. 1264–1270. Available at: https://doi.org/10.1109/ICICCS48265.2020.9121114.

Saleem, S., Aslam, M. and Rukh Shaukat, M. (2021) ‘A REVIEW AND EMPIRICAL COMPARISON OF UNIVARIATE OUTLIER DETECTION METHODS’, 37(4), pp. 447–462.

Santos, M.S. *et al.* (2018) ‘Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [Research Frontier]’, *IEEE Computational Intelligence Magazine*, 13(4), pp. 59–76. Available at: https://doi.org/10.1109/MCI.2018.2866730.

Singh, S. and Srivastava, R.K. (2020) ‘Understanding the intention to use mobile banking by existing online banking customers: an empirical study’, *Journal of Financial Services Marketing*, 25(3–4), pp. 86–96. Available at: https://doi.org/10.1057/S41264-020-00074-W/METRICS.

Smiti, A. (2020) ‘A critical overview of outlier detection methods’, *Computer Science Review*, 38, p. 100306. Available at: https://doi.org/10.1016/J.COSREV.2020.100306.

Sohony, I., Pratap, R. and Nambiar, U. (2018) ‘Ensemble learning for credit card fraud detection’, *ACM International Conference Proceeding Series*, pp. 289–294. Available at: https://doi.org/10.1145/3152494.3156815.

Stiglic, G. *et al.* (2020) ‘Interpretability of machine learning-based prediction models in healthcare’, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(5), p. e1379. Available at: https://doi.org/10.1002/WIDM.1379.

Taneja, S., Suri, B. and Kothari, C. (2019) ‘Application of Balancing Techniques with Ensemble Approach for Credit Card Fraud Detection’, *International Conference on Computing, Power and Communication Technologies (GUCON)* [Preprint].

Thorat, H., Thakur, S. and Yadav, A. (2020) ‘Categorization of Illegal Activities on Dark Web using Classification’, *International Research Journal of Engineering and Technology* [Preprint]. Available at: http://parazite.nn.fi/roguesci/ (Accessed: 4 May 2023).

Tudeal K.Y.D. (2022) ‘Internet Banking Development as A Means of Providing Efficient Financial Services in South Sudan’, *Specialusis Ugdymas*, 2(43), pp. 139–141.

UK Finance (2022) *ANNUAL FRAUD REPORT THE DEFINITIVE OVERVIEW OF PAYMENT INDUSTRY FRAUD IN 2021*, *UK FINANCE ANNUAL FRAUD REPORT*. Available at: https://www.ukfinance.org.uk/system/files/2022-06/Annual%20Fraud%20Report%202022\_FINAL\_.pdf (Accessed: 5 May 2023).

Varmedja, D. *et al.* (2019) ‘Credit Card Fraud Detection - Machine Learning methods’, *2019 18th International Symposium INFOTEH-JAHORINA, INFOTEH 2019 - Proceedings* [Preprint]. Available at: https://doi.org/10.1109/INFOTEH.2019.8717766.

Venkatanagendra, K. and Ussenaiah, M. (2019) ‘XGB CLASSIFICATION TECHNIQUE TO RESOLVE IMBALANCED HEART DISEASE DATA’, *IJRECE*, 7(1), pp. 406–411.

Van Vlasselaer, V. *et al.* (2015) ‘APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions’, *Decision Support Systems*, 75, pp. 38–48. Available at: https://doi.org/10.1016/J.DSS.2015.04.013.

Walker, M.L. *et al.* (2018) ‘An Improved Boxplot for Univariate Data’, *https://doi.org/10.1080/00031305.2018.1448891*, 72(4), pp. 348–353. Available at: https://doi.org/10.1080/00031305.2018.1448891.

Windasari, N.A. *et al.* (2022) ‘Digital-only banking experience: Insights from gen Y and gen Z’, *Journal of Innovation & Knowledge*, 7(2), p. 100170. Available at: https://doi.org/10.1016/J.JIK.2022.100170.

Wu, T.Y. and Wang, Y.T. (2021) ‘Locally Interpretable One-Class Anomaly Detection for Credit Card Fraud Detection’, *Proceedings - 2021 International Conference on Technologies and Applications of Artificial Intelligence, TAAI 2021*, pp. 25–30. Available at: https://doi.org/10.1109/TAAI54685.2021.00014.

Xiao, Z., Wang, L. and Du, J.Y. (2019) ‘Improving the performance of sentiment classification on imbalanced datasets with transfer learning’, *IEEE Access*, 7, pp. 28281–28290. Available at: https://doi.org/10.1109/ACCESS.2019.2892094.

Xu, T., Coco, G. and Neale, M. (2020) ‘A predictive model of recreational water quality based on adaptive synthetic sampling algorithms and machine learning’, *Water Research*, 177, p. 115788. Available at: https://doi.org/10.1016/J.WATRES.2020.115788.

Xuan, S. *et al.* (2018) ‘Random forest for credit card fraud detection’, *ICNSC 2018 - 15th IEEE International Conference on Networking, Sensing and Control*, pp. 1–6. Available at: https://doi.org/10.1109/ICNSC.2018.8361343.

Yang, L. and Shami, A. (2020) ‘On hyperparameter optimization of machine learning algorithms: Theory and practice’, *Neurocomputing*, 415, pp. 295–316. Available at: https://doi.org/10.1016/J.NEUCOM.2020.07.061.

Yash Chaushan (2018) *How Important Is Digital Banking*, *Tech Funnel*. Available at: https://www.techfunnel.com/fintech/how-important-is-digital-banking/ (Accessed: 3 May 2023).

Yee, O.S., Sagadevan, S. and Malim, N.H.A.H. (2018) ‘Credit Card Fraud Detection Using Machine Learning As Data Mining Technique’, *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 10(1–4), pp. 23–27. Available at: https://jtec.utem.edu.my/jtec/article/view/3571 (Accessed: 28 May 2023).

Zhang, X. *et al.* (2021) ‘HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture’, *Information Sciences*, 557, pp. 302–316. Available at: https://doi.org/10.1016/J.INS.2019.05.023.

Zhou, H., Chai, H. feng and Qiu, M. lin (2018) ‘Fraud detection within bankcard enrollment on mobile device based payment using machine learning’, *Frontiers of Information Technology and Electronic Engineering*, 19(12), pp. 1537–1545. Available at: https://doi.org/10.1631/FITEE.1800580/METRICS.