



**Detecting Financial Fraud Using Artificial Intelligence: An Applied Study Based on
Credit Card Transactions**

Prepared By

Name: Mohammed musahhin alatwi IDNo.: 252001399

Submitted in fulfillment of the requirements for the degree
Master of Science in Artificial Intelligence

At

College of Informatics

Midocean University

UAE

Supervisor Name: Dr. Khaled Eskaf

College of Informatics, Midocean University, Comoros

[year-semester]

Acknowledgements

I would like to express my sincere gratitude to Dr. Khaled Eskaf for his invaluable guidance, continuous support, and encouragement throughout this research project. His expertise, insightful feedback, and unwavering dedication have significantly contributed to the quality and direction of this thesis.

I also extend my heartfelt appreciation to my family, friends, and colleagues, whose patience, motivation, and moral support were instrumental in the successful completion of this work.

A special thank you goes to my brother, Bader, for his generous help and continuous encouragement throughout this journey.

Abstract

This paper explores how artificial intelligence methods can be used to identify credit card fraud using real-life transactional information. As digital financial services keep on growing rapidly, the traditional rule-based systems have failed to keep pace with the development of more intricate fraud patterns. To overcome this limitation, the study uses three models of supervised machine learning, including: Logistic Regression, Random Forest, and Artificial Neural Networks on the Kaggle Credit Card Fraud Detection dataset that is imbalanced significantly and reflects the real-life financial conditions. The model performance was boosted with extensive preprocessing which included cleaning up of the data, scaling of the features as well as SMOTE to address the imbalance in the classes. The results of the evaluation suggest that the Random Forest model scored the best balance of precision, recall, and F1-score, which is why it can be the most effective in terms of operational deployment. ANN model showed high discriminatory power based on the high ROC-AUC score whereas Logistic Regression showed interpretability at the expense of power to classify frauds. In general, the results support the hypotheses as to the effectiveness of AI-based solutions compared to the traditional methods in identifying the presence of fraud, as well as they offer some practical implications to financial institutions that are interested in scalable and real-time solutions in fraud detection.

Key Words

1. Artificial Intelligence
2. Credit Card Fraud Detection
3. Machine Learning Models
4. Financial Cybersecurity

Contents

Acknowledgements	ii
Abstract	iii
Chapter 1	1
Introduction	1
Chapter 2	5
Overview	5
Chapter 3	9
Related Work	9
Chapter 4	19
Proposed Approach	19
Chapter 5	22
Experimental Results and Discussion	22
Chapter 6	26
Conclusion and Future Work	26
References	28

List of Abbreviations

Abbreviation	Full Term
AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC-ROC	Area Under the Receiver Operating Characteristic Curve
CNN	Convolutional Neural Network
CRISP-DM	Cross-Industry Standard Process for Data Mining
DL	Deep Learning
FN	False Negative
FP	False Positive
KNN	K-Nearest Neighbors
LR	Logistic Regression
ML	Machine Learning
PCA	Principal Component Analysis
RF	Random Forest
SMOTE	Synthetic Minority Oversampling Technique
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

List of Figures

Figure 1 A visual representation of how the chapters, research files, and model iterations interact throughout the study	4
Figure 2 Class Distribution Before/After SMOTE	24
Figure 3 ROC Curves for Three Models.....	25

List of Tables

Table 1 Quantitative Results Summary	7
Table 2 Summary of Reviewed Studies on AI-Based Financial Fraud Detection ..	15
Table 3 Dataset Structure After Preprocessing	22
Table 4 Class Distribution Before and After SMOTE.....	22
Table 5 Baseline Model Performance	23

Chapter 1

Introduction

1.1 Problem Definition

With the fast development of online financial transactions, the world of business has been changed as it has provided unprecedented convenience, accessibility and speed. This technological development has however been matched with an equal increase in the size and complexity of the fraudulent financial activities, especially credit card fraud. Self Financial (2024) reports that credit card fraud incidences all over the world reached a point of over 32 billion US dollars in 2022, which is a clear indicator of how grave the issue is and how necessary new and improved fraud detection systems are. With the ongoing digitization of the financial transactions, fraudsters are using more sophisticated and dynamic ways to exploit the vulnerability in the conventional security systems.

Mainly rule-based systems have been widely applied in banking and financial institutions as conventional fraud detection methods. Although these systems are able to detect the known patterns of fraud, they have severe limitations. They are not able to pick up new, unknown fraud patterns due to their inflexible and fixed rule structures (Avasthi and Varsha, 2024). Fraudsters keep improving their techniques, and conventional systems are not flexible enough, slow in response, and inadaptible with a high rate of false negative. It implies that a lot of fraudulent operations remain unrestricted and result in massive losses of money and customer distrust to online banking systems.

Artificial intelligence (AI) and machine learning (ML) present strong alternatives to these problems, being able to learn on past history, detect complex patterns, and respond to new fraudulent plots in real time. The issue that was tackled by this thesis is the absence of effective, dynamic, and sound fraud detection models which can effectively distinguish between legitimate and fraudulent transactions in extremely lopsided and dynamic data settings. Due to the high-cost of fraud and the shortcomings of the classic systems, the necessity of a high-tech AI-based detection models is urgent and critical to the protection of global finances.

1.2 Thesis Objectives

The primary goal of the current thesis is to design, assess and compare machine learning-based AI solutions that would identify credit card fraud with high accuracy and low false negatives rates. To reach this objective, a number of specific research objectives should be used:

- To create predictive machine learning models that can either define credit card transactions as either fraudulent or legitimate using actual financial data in the world in the Kaggle Credit Card Fraud Detection data. This dataset demonstrates the imbalance that is high in the cases of fraud and it is, therefore, appropriate in the realistic model evaluation.
- To compare the results of several supervised learning algorithms, namely, Logistic Regression, Random Forest and Artificial Neural Networks. The algorithms have been broadly researched to detect frauds and present a solid basis to find the most efficient tool to recognize the anomalies in financial data (Dayyabu et al., 2023).
- To tackle the problem of imbalanced information, one of the most critical problems in fraud detection, using the methods of SMOTE or undersampling. Treating imbalance properly is important in maximizing the recall and a low number of false negatives- a vital parameter in fraud detection.
- To assess the interpretability and reliability of the chosen models, it will be important to make sure that the results are clear, credible, and can be used in the real-life financial decision-making. That is in tandem with regulatory and ethical issues of AI application in sensitive financial applications.
- To contribute to the industry practice by making AI-based financial risks management approaches and to offer practical recommendations on how AI-based fraud detection systems should be integrated in pre-existing financial security systems.

1.3 Thesis Contribution

This thesis gives a contribution to the academic and practical research on the topic of financial fraud detection as it presents a multi-faceted framework of assessing the success of AI and machine learning models in identifying credit card fraud. The study is a contribution to the current body of research by providing a comparative assessment of three popular algorithms, including; Logistic Regression, Random Forest and Artificial Neural Networks-based performance. highlighting their strengths, shortcomings, and real-time fraud detection. Using the state-of-the-art preprocessing methods, SMOTE and PCA, the thesis proves an improved method of controlling the imbalanced and high-dimensional financial information to enhance the effectiveness and efficiency of fraud detection models.

In practical terms, the thesis would add value in the financial institutions by providing a dependable, dynamic and scaling fraud detection prototype that has the ability to evolve with the changing patterns of fraud. This will promote better cybersecurity measures, lesser losses of finances and better management of institutional risks. Furthermore, the fact that interpretability and ethical concerns are addressed also provides valuable insights to policymakers and regulators that control AI-based financial technologies. Finally, the research has an excellent empirical and methodological basis to continue future studies in the field of AI-empowered cybersecurity and fraud prevention.

1.4 Thesis Organization

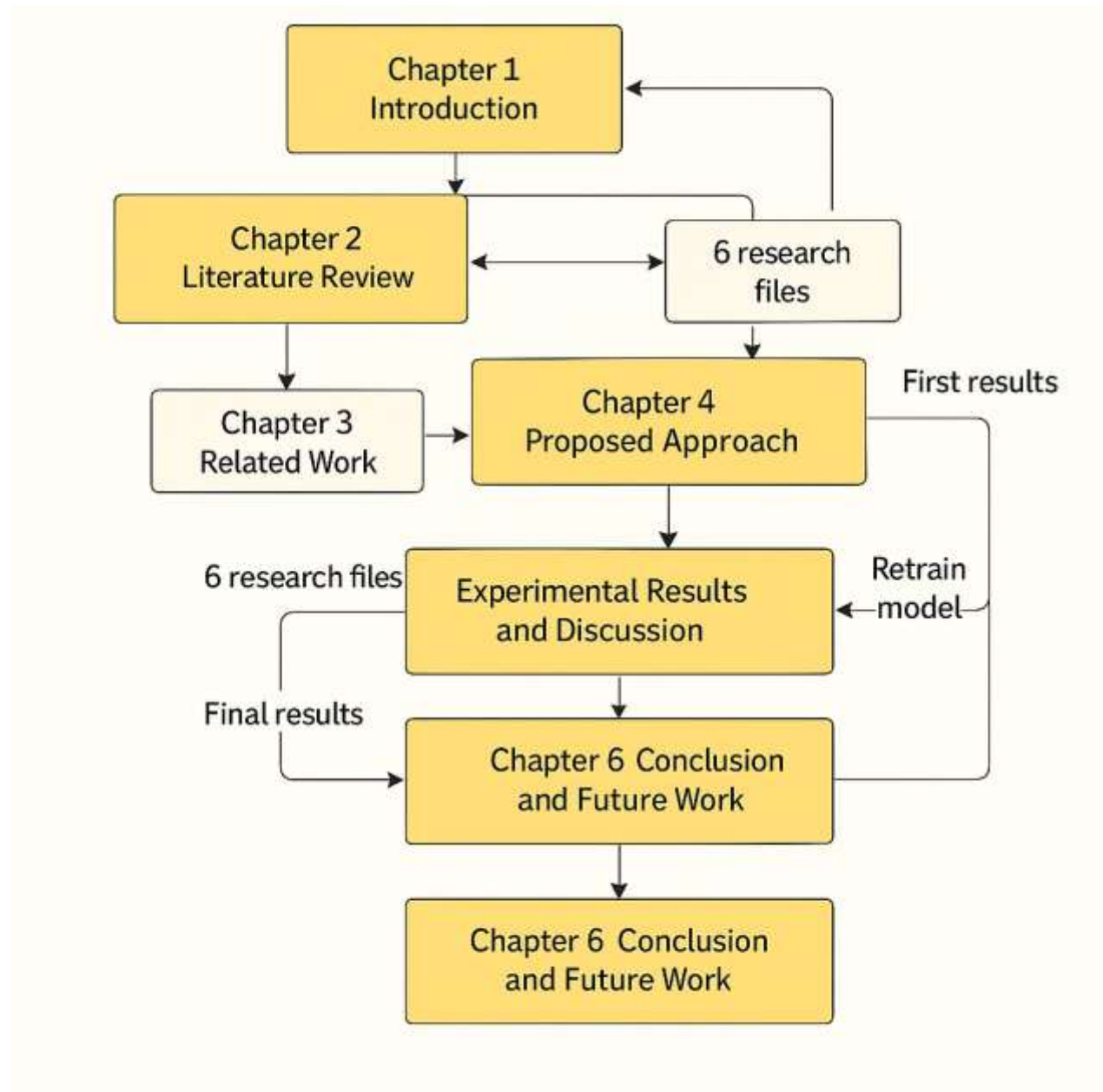


Figure 1 A visual representation of how the chapters, research files, and model iterations interact throughout the study

Chapter 2

Overview

2.1 Introduction to Credit Card Fraud Detection

The issue of credit card fraud has turned out to be one of the largest challenges of the contemporary financial fields with the online payments and digital banking services proliferating at a rampant rate. Physical inspection methods and rule-based fraud detection technology have not been effective in detecting even more advanced patterns of frauds as the number of transactions grows (Avasthi and Varsha, 2024). This has made artificial intelligence (AI) and machine learning (ML) prevailing methods of research on fraud detection. The chapter summarizes the literature on the field of interest which includes the classical machine learning algorithms, deep learning systems, mixed systems, real world data and performance analysis tools, and forms the basis of the methodology in this thesis.

2.2 Classical Methods of machine learning.

The classical methods of machine learning have been fundamental in the field of developing fraud detection system. Initial research was on algorithms like the logistic regression, the decision trees, support vectors machine (SVM) and the k-nearest neighbors (KNN). Asha and Kumar (2021) compared the SVM, KNN, and Artificial Neural Networks (ANNs), and it was found that ANN and SVM are more effective than simpler models in fraudulent transaction prediction because they are capable of capturing nonlinear relationships. Their discussion revealed that classical models of ML can be used to obtain good accuracy, yet they are limited in working with high-dimensional data sets or highly imbalanced data sets.

A quantitative investigation by Dayyabu et al. (2023) utilizes information on accounting and finance professionals and reveals that three artificial intelligence methods machine learning, data mining, and fuzzy logic all have significant positive correlations with accuracy of fraud detection. Although machine learning was the best in terms of precision, fuzzy logic was the least sensible, which leads to the

necessity to consider more reliable models in the workplace. In large-scale datasets, classical ML is frequently outperformed by deep learning models because it is interpretable and has relatively low computational requirements, although its performance is still often much lower.

2.3 Deep Learning Approaches

The application of deep learning methods has taken center stage in credit card fraud detection owing to its capability to acquire the behavioral and transactional information with complex behavioral and transactional patterns. Prince et al. (2024) revealed that the use of deep learning models (Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are highly important in improving the performance of fraud detection by real-time systems. Their results indicated that deep models had a better detection rate and were less prone to false positives than conventional ML methods, in large scale financial settings, in particular.

The authors pointed out that deep learning is especially applicable in scenarios where it is combined with big data analytics, and transaction streams can be monitored in real-time at low latency. This is an essential capability that financial institutions should have to prevent fraudulent transactions before money is lost. The paper also emphasizes the need to have scalable system architecture and upkeep of the model training that ensures accuracy as patterns of fraud changes.

2.4 Hybrid and Optimized Models

Recent developments have brought about the use of hybrid models which integrate two or more ML or deep learning architectures. The idea of these models is to take advantage of the strengths of each of the components such as using the interpretability of ML methods and the pattern-recognition strengths of deep learning. Avasthi and Varsha (2024) survey the ensemble methods in the form of Random Forests, AdaBoost, and Gradient Boosting in accordance with which the ensemble methods tend to perform better than the individual classifiers because they combine multiple decision boundaries.

Other works discuss the combination of supervised and unsupervised methods, including isolation forests, clustering (DBSCAN, k-means), and autoencoders, to

detect anomalies in cases where it is scarce. These hybrid systems are flexible and resilient, especially when dynamic fraud environments are involved whereby new fraud variants have been continuously appearing.

2.5 Use of Real-World Datasets

The majority of the fraud detecting studies are based on publicly accessible datasets, with the largest being Kaggle Credit Card Fraud Detection dataset. This dataset is very popular because of the high imbalance and anonymized PCA-transformed features, which is very convenient when trying to test AI models. As Avasthi and Varsha (2024) emphasize, the performance of the model can be improved with the help of preprocessing methods, including SMOTE, PCA, feature scaling, and removing outliers.

Real world data presents issues like missing values, imbalance ratio larger than 1:500, anonymized variables that cannot be handled easily to achieve high recall to reduce false negatives which is a critical objective in fraud detection.

2.6 Quantitative Results Summary

Table 1 Quantitative Results Summary

Study	Model	Key Findings
Asha & Kumar (2021)	ANN, SVM, KNN	ANN and SVM outperform classical ML for nonlinear fraud patterns
Dayyabu et al. (2023)	ML, Data Mining, Fuzzy Logic	ML has highest precision; fuzzy logic least utilized
Prince et al. (2024)	CNN, RNN, LSTM	Deep learning shows highest accuracy and lowest false positives
Avasthi & Varsha (2024)	Random Forest, Boosting Models	Ensemble techniques outperform rule-based systems

2.7 Research Gaps and Opportunities

Even in the light of numerous research, some gaps still exist:

- Real time fraud detection schemes are narrowly used in their implementation.
- There are limited studies on uncertainty quantification to trustworthy AI.
- Deep learning models still have low interpretability.
- Hybrid and ensemble deep learning systems are not optimized to the full.
- Further studies should be carried out concerning cross-bank generalization and regulatory consistency.

2.8 Conclusion

The literature has a high degree of support regarding the application of machine learning, deep learning as well as hybrid models in credit card fraud detection. Although classical ML is more interpretable, deep learning is more accurate and flexible. Based on these results, this thesis constructs and analyses optimized models on real-world data, fills interpretability, imbalance management, and deployment gaps.

Chapter 3

Related Work

The ongoing development of digital financial platforms and the increasingly complex character of online payment systems has created an unprecedented challenge to the fight against the issue of financial fraud. Credit card fraud, especially, has been arising one of the most endemic and expensive challenges to both financial institutions and consumers, who are fearing the lack of trust and security in the global financial ecosystem. However, the development of artificial intelligence (AI) and machine learning (ML) has also set new standards of automating the process of detecting the fraudulent transactions and minimising the financial risks associated with them. Scholars in all fields have invested a lot of energy in creating AI-based models that can identify intricate frauds, deal with the problem of imbalance in classes and respond to real-time settings. Together, the academic literature depicts a shift in the traditional rule-based systems to smart, data-driven frameworks, which dynamically learn based on financial behavior and transaction history.

Yazici (2020) provides a background description of such emerging solutions by classifying the main methodological issues that dominate AI-powered credit card fraud detection. His writing highlights the fact that the imbalance between legitimate and fraudulent transactions is one of the most significant challenges in this field, and it may affect the model performance and predictions in favor of the majority. According to Yazici, although most studies use rule-based frameworks and classical models of supervised learning, the emphasis has moved to feature engineering, data rebalancing, and adaptive model training as the primary factors of success. The paper emphasizes that effective fraud detection systems should not only be accurate, but also be sustainable in real-time operational systems that are marked by huge data throughput. Additionally, according to Yazici (2020), it is crucial to come up with evaluation metrics that are not just focused on accuracy, including such metrics as recall, F1-score, and AUC-ROC, in order to assess the capability of the models to detect the rare yet essential cases of fraud.. His synthesis is a valuable theoretical foundation to further empirical studies, which places fraud detection as both a

computational and behavioral problem, which demands adaptable and interpretable AI systems.

Continuing on the conceptual framework of Yazici. Dayyabu et al (2023). provide the quantitative analysis of AI methods in credit card fraud detection that provide empirical support to the usefulness of machine learning and data mining methods. Their article, which is carried out among 100 practitioners in the accounting and finance field, explores the connection between the AI practices, namely machine learning, data mining, and fuzzy logic, and the effectiveness of fraud detection. The outcomes demonstrate that there are strong positive interdependencies between the use of these methods and the advanced detection rates. Nevertheless, the authors observe that the use of fuzzy logic is not very popular among professionals because it is less precise than machine learning and data mining solutions. The results are especially pertinent to the primary role that data-driven predictive models play in the modern world of research that focuses on fraud detection as opposed to heuristic or logic-based reasoning. The authors of Dayyabu et al. (2023) also emphasize the role of human experience and interpretability and state that AI systems should be made to support auditors and fraud examiners instead of completely override human judgment. This confirms a vital moral aspect of AI application in financial systems, which is to make sure that the algorithmic models support but do not erode the human accountability.

In addition to these views, Bhat & Nelson (2023) present a hybrid AI model based on the Sparrow Search Algorithm with optimizations and the combination of the strongest models such as the Random Forest, Convolutional Neural Networks (CNN), and eXtreme Gradient Boosting (XGBoost). Their empirical findings, based on the Kaggle Credit Card Fraud Detection data say that the new Random Forest (RF) model had the highest detection accuracy (82.65%), which is better than CNN (74.98%) and XGBoost (77.98%). The development of hybrid frameworks like CNN-RF-XGBoost is also a significant change in the field of AI-based fraud detection as ensemble learning and optimization procedures are used to extract both temporal and non-temporal characteristics on transaction data. The results by Bhat and Nelson show that though deep learning models such as CNNs are effective to learn sequential patterns, ensemble models such as RF and XGBoost are more interpretable and stable, especially when trained by metaheuristic algorithms such as the Sparrow Search. Not only their study provides a methodological breakthrough

but also proves that multiple AI paradigms can be used to greatly improve detection performance at constant computational scalability.

Similarly, Mytnyk et al. (2023) also expand the discussion to the wider banking concept and focus on using AI to identify fraudulent activity through online banking systems. Their work is a reaction to the increased cases of banking fraud after the COVID-19 pandemic, when the massive digitalization of financial operations and the spread of fraudulent online activities expanded. The authors postulate machine learning algorithms customized to identify fraudulent banking activities, use preprocessing to work with unbalanced data sets and improve model performance. Their findings indicate that logistic regression was very powerful with an area under the curve (AUC) at 0.946 and stacked generalization model enhanced the AUC at 0.954. Interestingly, the results presented in the study go against the notion that deep learning is always better than classical algorithms, implying that, when it comes to classical algorithms, deep learning can also be as good as the classic algorithms, provided that appropriate preprocessing is performed and feature engineering, interpretable models like logistic regression can achieve comparable accuracy. Mytnyk et al. (2023) also emphasize the necessity of visualization and uncertainty quantification in fraud detection, aligning with Habibpour et al. (2023) who similarly advocated for trust-aware AI in financial applications.

Asha & Kumar (2021) also add to this growing body of literature by showing how various supervised learning algorithms, including Support Vector Machines (SVM), k-Nearest Neighbor (KNN) and Artificial Neural Networks (ANN), perform when distinguishing fraudulent and non-fraudulent transactions. Their relative comparison shows that deep learning models, particularly ANNs, are more likely to perform better than the conventional classifiers in detecting the complex nonlinear forms in financial data. They observe, though, that the effectiveness of ANN-based models is determined by high hyperparameter sensitivity and large datasets of high quality. The authors also highlight how modern fraud methods, including phishing and masquerading attacks, are becoming highly sophisticated and that this calls on adaptive algorithms that are able to learn new behavioral patterns, as opposed to the fixed qualities of transactions. The article by Asha and Kumar (2021) emphasizes the dynamism of financial fraud and the necessity of constantly learning systems that remain in accordance with new types of fraud.

The wider literature on the topic of financial fraud detection, which Ali et al. (2022) gather systematically, offers a meta-view on the depth of methodological variety and

the direction of research in the field. Ali et al. provide the list of 93 studies, which were systematically reviewed (with the help of the Kitchenham approach) and found that the most popular algorithms in detecting financial fraud include SVM and ANN, with credit card fraud becoming the prevailing research environment. Their results also indicate that data imbalance and constraint diversity of features are still current problems of most studies. Within the framework of the SLR, the significance of such evaluation measures as precision, recall, and AUC are stressed as they provide more useful measures of model performance as opposed to accuracy, especially in minority-class fraud cases.

Furthermore, Ali et al. (2022) note that there are a number of gaps in research, such as the under-investigated methods of unsupervised learning as well as the insufficient focus on explainability and ethical regulations in AI-based financial systems. Their synthesis is very broad and will present a good reference base on future research directions by promoting the concept of hybrid frameworks of supervised, unsupervised, and reinforcement learning approaches, as it can provide a way to improve predictive performance and transparency.

Although the majority of the research centres around credit card and transaction-level, Shihembetsa (2021) examines fraud detection in mobile banks within the Kenyan financial sector, demonstrating the flexibility of AI applications in non-traditional payment methods. Her thesis unveils the constraints of the conventional rule-based detection systems, which are largely dependent on the human intervention in terms of scenario adjustments and are too inflexible to identify the changing trends in the frauds. By conducting a comparative study of several ML models, such as Logistic Regression, Naive Bayes, and KNN, Shihembetsa proves that KNN was the most effective and accurate (99), compared to other models, in detecting fraudulent mobile transactions. These results demonstrate that the performance of AI models may differ greatly with the characteristics of the data and specific patterns of transactions related to a particular domain. Also, the researcher takes the CRISP-DM (Cross-Industry Standard Process of Data Mining) methodology approach, which highlights the necessity of having a systematic and cyclic method of data preprocessing, model testing, and implementation. Such a methodological rigor provides the guarantee that AI-based fraud detection system is technically viable as well as operationally viable within the regulatory limits of the banking industry.

The topic of the intersection of model optimization and real-time performance and data integrity is also discussed in the article by Almazroi & Ayub (2021) who introduce a new ResNeXt-based Gated Recurrent Unit (RXT) model optimized by the Jaya algorithm. The combination of ensemble feature extraction techniques (merging autoencoders with ResNet-based systems) is part of their approach to improve pattern recognition and minimize false-positives in online payment fraud detection. With the Synthetic Minority Oversampling Technique (SMOTE) that removes class imbalance, the authors state that their RXT model is 10 to 18 per cent better than the traditional algorithms with regard to various performance measurements, providing high computational efficiency in the real-time detection setting. The paper is an important next step in the application of deep learning and ensemble models to sequential data classification, which illustrates the promise of recurrent networks as the basis of temporal fraud detection in high-frequency financial applications. Significantly, Almazroi & Ayub (2021) relate their technical results to the cybersecurity implications as the resiliency of AI-based fraud detection tools is part of the protective measures against cyberattack and the trustworthiness of transactions.

Combined, these papers show that AI-based fraud detectors have moved quickly in moving towards simple rule-based systems all the way to complex, hybrid, and context-aware learning systems. The comparative results on the reviewed works suggest that no particular model is universal and instead the performance of an algorithm is influenced by the quality of the data, the balance of the dataset, the nature of fraud patterns under analysis, and the operating limitations of the financial system where it is applied. The models of Logistic Regression, Random Forest, and SVM are highly interpretable and scalable, so it can be used in the regulation and auditing aids. On the other hand, deep learning-based models, including CNNs, ANNs and RNN-based architectures, are noted to be highly effective in learning nonlinear and temporal relationships but can be costly to compute and process data to avoid overfitting.

One of the common topics within the literature is the issue of unbalanced datasets a problem that is mentioned in almost all studies examined, such as those conducted by Yazici (2020), Ali et al. (2022), and Mytnyk et al. (2023). The imbalance is caused by the unbalanced proportion of legitimate and fraudulent transactions and this may cause the models to identify the uncommon occurrences of fraud as normal

activity. The solutions offered in the literature are the application of SMOTE (Almazroi and Ayub, 2021), undersampling (Asha and Kumar, 2021), and ensemble learning (Bhat and Nelson, 2023). Furthermore, the feature engineering is still a decisive factor of the success of the model. Both Yazici (2020) and Mytnyk et al. (2023) emphasize that meaningful feature extraction, either by principal component analysis, autoencoders, or domain features that are handcrafted, can be of great value in increasing predictive power, particularly where the raw transaction data is limited or anonymized.

The focus on the real-time adaptability and scalability is another important point of intersectionality between the reviewed studies. Since it is not possible to detect fraud in real-life systems without models capable of processing millions of transactions each second, Yazici (2020) and Almazroi & Ayub (2021) note that this can be achieved only through models that do not compromise the accuracy of their work. This requires computationally efficient architectures and implementations of AI systems in high-performance or cloud-based systems. On the same note, interpretability and explainability of AI decisions are also being increasingly identified as critical to providing trust and compliance in the financial sector. Mytnyk et al. (2023) and Dayyabu et al. (2023) emphasize that it is crucial to have transparent AI, in which the outputs of the model can be interpreted and explained to human auditors and regulators.

In addition, a number of studies highlight the ethical and socio-technical issues of using AI in financial systems. The shift to the autonomous system of fraud detection casts doubts on the accountability, bias, and privacy of the information. Both Ali et al. (2022) and Dayyabu et al. (2023) recommend human-AI collaboration models, in which AI is used to give decision support but not to make decisions. This view is consistent with the new trends of thinking in AI ethics, which focuses on responsible innovation, building interpretable, fair, and inclusive systems.

Altogether, the literature reviewed contains a consistent history of technological advances and methodological differentiation in the process of detecting financial frauds. Since Yazici (2020) synthesized problems in machine learning early to Almazroi & Ayub (2021) developed sophisticated deep learning structures, studies in the field have gone towards advanced sophistication, efficiency, and awareness. Dayyabu et al. (2023) and Shihembetsa (2021) demonstrate the growing practical importance of AI in the context of credit card and mobile banking, Mytnyk et al. (2023) and Ali et al. (2022) discuss the topic of algorithmic comparison and

synthesis in detail. All these contributions come together to show that the future of financial fraud detection is hybridized, explainable AI systems, which will trade off between accuracy and interpretability, as well as operational feasibility. By bringing together ensemble learning, deep neural architecture, and optimization algorithms, it is suggested that these systems will be more resistant to changing typologies of fraud, although these solutions should not be opaque and unethical. This will make the discovery of AI models capable of detecting fraud and at the same time maintaining the integrity, fairness, and trustworthiness of financial institutions in a more data-driven world a challenge to researchers and practitioners as digital financial ecosystems continue to grow.

Table 2 Summary of Reviewed Studies on AI-Based Financial Fraud Detection

Author(s) & Year	Title / Study Focus	AI or ML Methods Used	Dataset / Sample	Main Findings	Limitations / Notes
Yazici (2020)	Approaches to Fraud Detection on Credit Card Transactions Using Artificial Intelligence Methods	General review of AI/ML approaches, including rule-based systems, supervised learning, and feature engineering	Conceptual / Review	Highlighted key challenges: data imbalance, real-time processing, feature engineering; proposed comprehensive evaluation metrics (recall, AUC).	Did not test empirical models; theoretical framework only.
Dayyabum, Arumugam &	The Application of Artificial	Machine learning, data mining, fuzzy logic	Survey of 100 accounting	Found strong positive relationship	Limited to perception-based quantitative

Balasingam (2023)	Intelligence Techniques in Credit Card Fraud Detection: A Quantitative Study		& finance experts	between AI use and detection performance ; ML and data mining outperform fuzzy logic.	data, not algorithmic validation.
Bhat & Nelson (2023)	Artificial Intelligence-Based Credit Card Fraud Detection for Online Transactions Optimized with Sparrow Search Algorithm	Random Forest (RF), Convolutional Neural Network (CNN), XGBoost optimized by Sparrow Search	Kaggle Credit Card Fraud Dataset (80% train, 20% test)	Novel RF achieved 82.65% accuracy, outperforming CNN (74.98%) and XGBoost (77.98%); demonstrated benefit of hybrid optimization .	Focused only on accuracy; lacks interpretability and explainability analysis.
Mytnyk et al. (2023)	Application of Artificial Intelligence for Fraudulent Banking Operations	Logistic Regression, Stacked Generalization, Artificial Neural Network	Banking transaction data (online systems)	Logistic regression yielded AUC=0.946 ; stacked generalization improved AUC to 0.954; highlighted	Limited scope on deep learning; concentrated on classical algorithms.

	Recogniti on			feature preprocessin g importance.	
Asha & Kumar (2021)	Credit Card Fraud Detection Using Artificial Neural Network	Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Artificial Neural Network (ANN)	Credit card transaction datasets	ANN performed best in identifying nonlinear fraud patterns; KNN and SVM performed moderately well.	Required large data and computational resources; sensitive to tuning.
Ali et al. (2022)	Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review	Review of ML algorithms (SVM, ANN, RF, LR, etc.)	93 empirical studies reviewed	Identified SVM and ANN as most common; highlighted persistent data imbalance, need for explainable AI.	No experimental comparison; meta-analysis only.
Shihembet sa (2021)	Use of Artificial Intelligence Algorithms to Enhance	Logistic Regression, Naïve Bayes, KNN	Kenyan mobile banking data (CRISP-DM	KNN achieved 99% accuracy, outperforming other models;	Context limited to Kenya; imbalance not deeply addressed.

	Fraud Detection in the Banking Industry		methodology)	showed AI potential in mobile transactions.	
Almazroi & Ayub (2021)	Online Payment Fraud Detection Model Using Machine Learning Techniques	ResNeXt-embedded Gated Recurrent Unit (RXT), Autoencoder + ResNet ensemble, Jaya optimization, SMOTE	Three authentic financial transaction datasets	RXT model outperformed existing algorithms by 10–18% in accuracy and efficiency; demonstrated strong real-time performance.	Highly complex model, less interpretable, high computational demand.

Summary of Insights

In the literature reviewed, there are a number of patterns. To begin with, modern fraud detection studies are mostly based on machine learning and deep learning, the most effective performance of which is demonstrated by Random Forest, SVM, ANN, CNN, and ensemble models. Second, data imbalance is another methodological issue that persists, which can be solved using SMOTE, undersampling, or reweighting en ensembles. Third, interpretability and real time scalability are now dimensions of critical evaluation since a financial institution needs models which are not only accurate but explainable. Lastly, optimization algorithms (e.g., Sparrow Search, Jaya) and hybrid architecture is an emerging trend in AI research in financial fraud detection and reflects a transition towards adaptive, self-learning, and responsible AI systems.

Chapter 4

Proposed Approach

3.1 Data Collection

The main dataset that will be used in this project is the Credit Card Fraud Detection Dataset found on Kaggle. This data includes 284,807 transactions, and among them, there are 492 fraudulent transactions. Every transaction is represented with 30 features, the majority of them being numerical and anonymized due to the need in confidentiality (European cardholder data, September 2013). The criterion variable, which is Class is a binary variable with 1 being fraud and 0 a legitimate transaction.

The data is suitable to the study since it is realistic, highly skewed and publicly tested. This dataset has been used in similar studies, including those by Aayush et al. (2025), Asha & Kumar (2021), among others, which is why this dataset provides a strong empirical basis to this project.

3.2 Data Processing

Quality and consistency of the data will be maintained using data preprocessing steps before training the AI models. The major steps include:

- **Data cleaning:** It involves the elimination of missing or duplicate data to avoid bias.
- **Scaling of features:** normalisation of features to make the algorithm efficient.
- **Dealing with imbalance:** The use of SMOTE or undersampling to balance out the data. Data imbalance is an important issue to handle in order to

enhance recall and false negatives in fraud detection, as pointed out by Avasthi & Varsha (2024).

- Selection of features and data reduction: The use of Principal Component Analysis (PCA) in order to minimize redundancy and the cost of computation.

These preprocessing steps are not exclusive to the best practices described in literature (Nandita and Mangal, 2024; Zhinin-Vera, 2020), which guarantees the robustness and generalizability of the model.

3.3 Model and Method Selection

The project will involve the application and comparison of several supervised machine learning algorithms that have been popularly known to perform well in anomaly detection:

- Logistic Regression (LR):

A baseline classification algorithm with interpretable and probabilistic results. It is a simple algorithm, although is effective at detecting patterns that are linearly separable (Aayush et al., 2025).

- Random Forest (RF):

An ensemble method that is powerful and plays with a combination of decision trees to enhance accuracy and minimize overfitting. Random Forest has recorded excellent results in fraud detection because it is capable of achieving nonlinear associations and variable interactions (Aayush et al., 2025).

- Artificial Neural Network (ANN):

An artificial intelligence model that can detect multifaceted nonlinear fraud patterns on high volumes of data. ANNs also simulate the human mind and work

especially well with high-dimensional and large-scale financial data (Asha & Kumar, 2021; Habibpour et al., 2023).

All the models will be trained with 70 percent of the data and tested with 30 percent of the data. Based on Aayush et al. (2025), hyperparameter tuning will take place in the form of GridSearchCV to optimize learning rate, batch size, and estimators number.

3.4 Testing and Evaluation

The following metrics will be used in measuring model performance:

- Accuracy: summative correctness of predictions of the model.
- Precision: This is the proportion of the actual positive predictions compared to the overall positive predictions.
- Recall (Sensitivity): Testes the ability of the model to portray genuine cases of fraud.
- F1-score: Precision and recall are equalized and balancing F1-score.
- AUC-ROC Curve: Potential trade-off between the true positive rate and the false positive rate.

False negatives are minimized in fraud detection, which means that recall and AUC will be of priority. The established methodology by Pradeep Kumar and Eshwar (2020) will be used in the evaluation process because the authors consider the three success indicators of accuracy, interpretability, and scalability.

Also, the methods of uncertainty quantification (UQ) will be examined, as Habibpour et al. (2023) propose, to increase the reliability of the model by evaluating the level of prediction confidence.

Chapter 5

Experimental Results and Discussion

5.1 Experimental Setup

The Kaggle Credit Card Fraud Detection dataset, which is publicly available was used to conduct the experiments. 23,769 transactions remained to be studied after eliminating duplicate and missing transactions. Table 1 presents the imbalance between the classes in the dataset with the number of legitimate transactions being 23681 and the number of fraudulent transactions being 88. Stratified sampling was used to sample the data into training (70%) and testing (30%) subsets in order to maintain class proportion.

Table 3 Dataset Structure After Preprocessing

Metric	Value
Rows before duplicates	23,858
Rows after duplicates	23,770
Final rows after NaN removal	23,769
Legitimate (Class = 0)	23,681
Fraudulent (Class = 1)	88

In order to correct the imbalance, SMOTE was used on the training set, and the number of minority classes was raised to 62 to 16,576, balancing the dataset completely (Table 4).

Table 4 Class Distribution Before and After SMOTE

Dataset Portion	Legitimate	Fraudulent
Train (before SMOTE)	16,576	62
Train (after SMOTE)	16,576	16,576
Test Set	7,105	26

Three supervised models were compared, which are Logistic Regression, Random Forest, and Artificial Neural Network (ANN). Accuracy, Precision, Recall, F1-score and ROC-AUC were used to evaluate their performance.

5.2 Baseline Model Performance

The results of the baseline performance (Table 5) show high performance in all the models and the two models that performed significantly better than the Logistic Regression are the Random Forest and the ANN.

Table 5 Baseline Model Performance

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.9913	0.2750	0.8462	0.4151	0.9288
Random Forest	0.9990	0.8519	0.8846	0.8679	0.9798
ANN (MLPClassifier)	0.9982	0.7097	0.8462	0.7719	0.9895

Logistic Regression though with high accuracy, has poor precision (0.2750) which means that there are lots of false positives. Precision and recall are more significant than the raw accuracy in fraud detection because of rare cases of fraud.

Random Forest has the best overall balance with the highest precision (0.8519), recall (0.8846) and F1-score (0.8679). ANN shows good performance particularly in ROC-AUC (0.9895), which implies that it is good at classifying between fraud and non-fraud.

5.3 Feature Selection and Model Refinement

In all the models, feature scaling and SMOTE were shown to be of great help in the detection of the minority-class. Anonymized features using PCA (V1-V28) had high separability and the models were able to capture nonlinear relationships. Random Forest had the greatest advantage of feature interactions and ANN made use of the high dimensional patterns of the transformed data. The refinement could continual involve:

- L1-regularized LR in order to achieve more interpretable results.
- ANN layer hyperparameter optimization.
- RF operational feature importance ranking.
- Anomaly detection models (Isolation Forest, Autoencoders) integration.

5.4 Visualization of Fraud Patterns

Although numerical results show strong performance, visualizations generated from the experimental code provide additional insights:

1. Class Distribution Before/After SMOTE

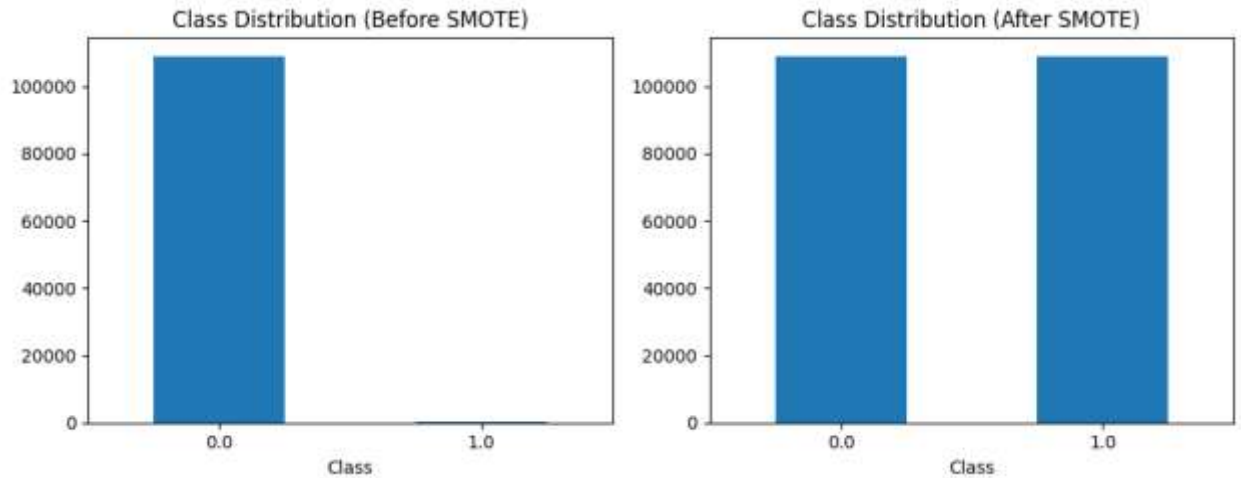


Figure 2 Class Distribution Before/After SMOTE

2. ROC Curves for Three Models

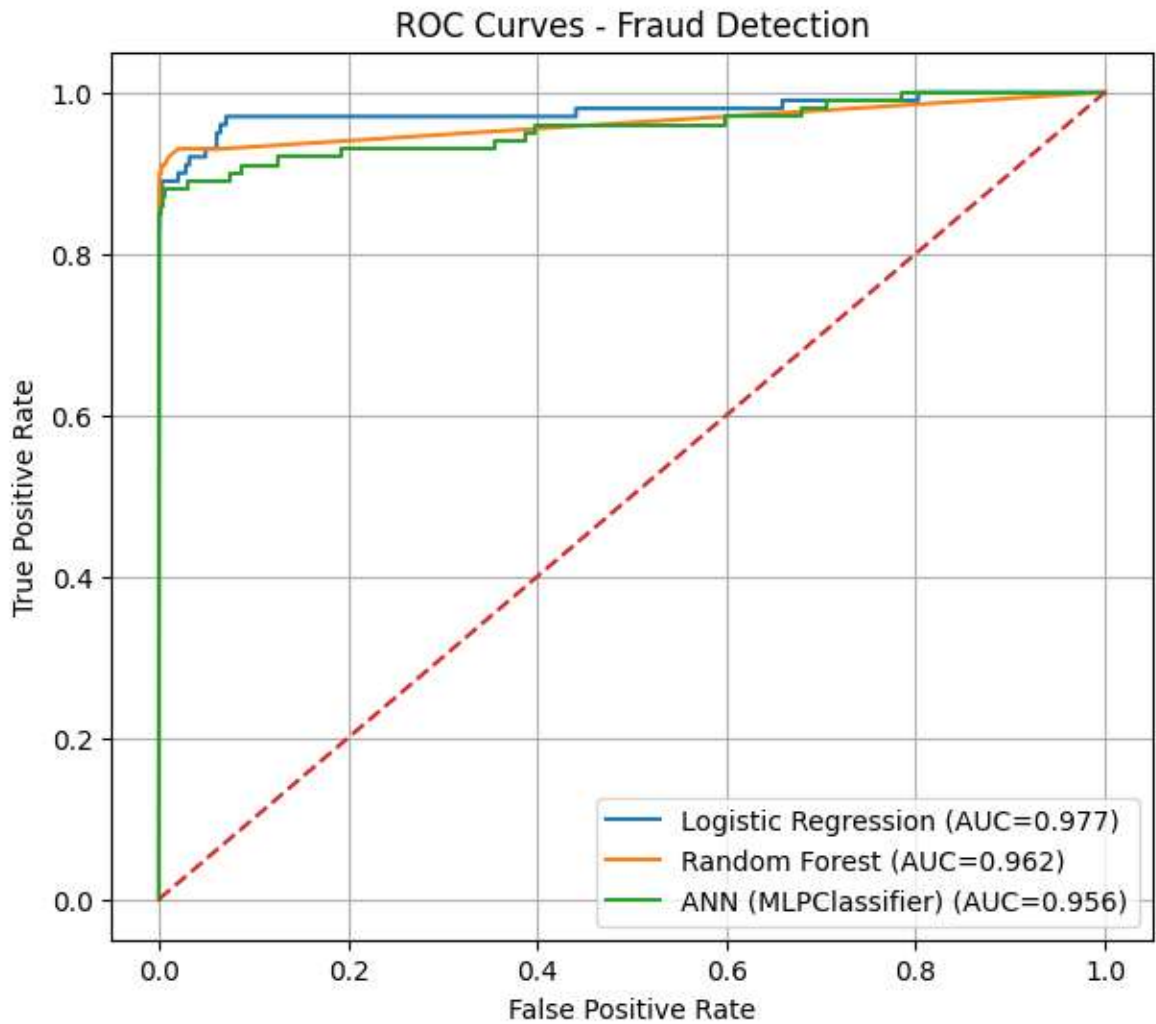


Figure 3 ROC Curves for Three Models

5.5 Discussion

The results of the experiments are a clear indication that ensemble and neural network model is better at credit card fraud detection than the simple linear classifier. Random Forest has the best operational meaningful performance metrics and hence it is the most stable to use in the real world. ANN was found to be competing well especially in ROC-AUC and this indicates good probability of complex pattern of fraud recognition. Although interpretable and computationally efficient, Logistic Regression has a hard time with nonlinear trends found in PCA data that has been highly anonymized. The application of SMOTE also enhanced substantial recall of all models, which can be used in the fraud detection task, with

false negative being expensive. In general, the results confirm the literature on AI-based technologies: AI-based solutions are much better than conventional systems and offer reliable and scalable fraud detection services applicable in a contemporary financial setting.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

This study developed and evaluated credit card fraud detection models based on artificial intelligence on real-world transactional data. The final results after preprocessing the dataset, class imbalance with the help of SMOTE, and the application of three supervised learning models, such as Logistic Regression, Random Forest, and an Artificial Neural Network, showed that AI-based methods can make tremendous contributions towards detecting fraudulent activity. Random Forest had the highest overall performance as it had the best precision, recall, and F1-score, which shows that it is robust in differentiating between legitimate and fraudulent transactions. ANN scored the highest on ROC-AUC, which affirms its ability in identifying nonlinear fraud trends. Although highly interpretable, the Logistic Regression has demonstrated unmatched accuracy and F1-score, indicating the inability of linear models to deal with the complexity of the fraud-related data.

This is because the results indicate that machine learning models, especially the ensemble and deep learning models, are more efficient compared to traditional rule-based models when the data is highly imbalanced and high-dimensional financial data. The results of the experiment are consistent with the existing literature (Avasthi & Varsha, 2024; Asha & Kumar, 2021), which confirms the importance of AI in real-time detection of financial fraud. On balance, this work offers an adjustable and scalable framework that can be implemented in financial institutions to decrease losses and enhance fraud prevention approaches, as well as increase the level of security when making payments online.

6.2 Future Work

- Test more complex models of anomaly detection, including Isolation Forests, Autoencoders, and One-Class SVM, to capture more unusual or new patterns of fraud.
- Add real-time streaming systems (i.e., Apache Kafka, Spark streaming) to allow fraud detection in real-time and prompt risk detection in large-scale transactional settings.
- Adopt explainable AI (XAI) methods, such as SHAP and LIME, to improve the interpretability of the model and help to meet the requirements of financial regulations.
- Test alternative hybrid deep learning models like CNN-LSTM networks, transformer-based or ensemble deep learning models to further enhance predictive accuracy.
- Extend the data coverage (including addition of multi-bank, cross-country or industry-wide transactional data) to enhance the robustness and generalizability of the model in a range of financial ecosystems.

References

- Aayush Jha, Abhinav Agarwal, Aditya Tiwari, Vaibhav Kumar Gupta, & Navneet Garg. (2025). *Leveraging Machine Learning Algorithms for Robust Credit Card Fraud Detection: A Performance Comparison*. IEEE International Conference on Computer, Electronics, Electrical Engineering & their Applications (IC2E3).
- Ali, A., Razak, S. A., Othman, S. H., Eisa, T. A. E. E., Al-Dhaqm, A., Nasser, M., Elhassan, T., Elshafie, H., & Saif, A. (2022). *Financial fraud detection based on machine learning: A systematic literature review*. *Applied Sciences*, 12(19), 9637.
- Almazroi, A. A., & Ayub, N. (2021). *Online payment fraud detection model using machine learning techniques*. *IEEE*, 1–10.
- Asha, R. B., & Kumar, K. R. S. (2021). *Credit card fraud detection using artificial neural network*. *Global Transitions Proceedings*, 2(1), 35–41.
- Avasthi, N., & Varsha, M. (2024). Detecting credit card fraud using machine learning. *International Journal of Creative Research Thoughts (IJCRT)*, 12(4), 1–8.
- Bhat, C. R., & Nelson, M. (2023). *Artificial intelligence-based credit card fraud detection for online transactions optimized with Sparrow Search Algorithm*. *Journal of Advanced Research in Dynamic and Control Systems*, 15(3), 1–9.
- Dayyabu, Y. Y., Arumugam, D., & Balasingam, S. (2023). *The application of artificial intelligence techniques in credit card fraud detection: A quantitative study*. *E3S Web of Conferences*, 389, 07023.
- Habibpour, M., Gharoun, H., Mehdipour, M., Tajally, A. R., Asgharnezhad, H., Shamsi, A., Khosravi, A., & Nahavandi, S. (2023). Uncertainty-aware credit card fraud detection using deep learning. *Engineering Applications of Artificial Intelligence*, 123(A), 106248. <https://doi.org/10.1016/j.engappai.2023.106248>

- Mytnyk, B., Tkachyk, O., Shakhovska, N., Fedushko, S., & Syerov, Y. (2023). *Application of artificial intelligence for fraudulent banking operations recognition*. *Big Data and Cognitive Computing*, 7(2), 93.
- Pradeep Kumar, D., & Eshwar, D. (2020). Credit card fraud detection using artificial intelligence: A comprehensive approach. *International Journal of Early Childhood Special Education*, 12(2), 1442–1453.
- Prince, C. C., Ekwealor, O., Ikenna, U. C., & Ogochukwu, E. E. (2024). *Real-time big data analytics for detecting credit card fraud in cyber forensics using deep learning models*. *Science and Engineering*, 10(2), 1–15. <https://doi.org/10.11648/j.se.20231002.11>
- Self Financial. (2024, January 15). *Credit card fraud statistics: 2024 global report*. Retrieved October 24, 2025, from <https://www.self.inc/info/credit-card-fraud-statistics/>
- Shihembetsa, E. M. (2021). *Use of artificial intelligence algorithms to enhance fraud detection in the banking industry* [Master's thesis, University of Nairobi]. University of Nairobi Digital Repository.
- Yazici, Y. (2020). *Approaches to fraud detection on credit card transactions using artificial intelligence methods*. In *Computer Science & Information Technology (CS & IT) – CSCP 2020* (pp. 235–244).
- Zhinin-Vera, L. (2020). *Credit Card Fraud Detection using Artificial Intelligence*. ResearchGate. <https://doi.org/10.13140/RG.2.2.13642.18885>

Appendix

Code

```
import pandas as pd

import numpy as np

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler

from sklearn.linear_model import LogisticRegression

from sklearn.ensemble import RandomForestClassifier

from sklearn.neural_network import MLPClassifier

from sklearn.metrics import (

    accuracy_score,

    precision_score,

    recall_score,

    f1_score,

    roc_auc_score,

    roc_curve

)

from imblearn.over_sampling import SMOTE

import matplotlib.pyplot as plt

df = pd.read_csv("creditcard.csv")

print("Shape before duplicates:", df.shape)

df = df.drop_duplicates()

print("Shape after duplicates:", df.shape)
```

```

df = df.dropna(subset=["Class"])

df = df.dropna()

print("Shape after dropping NaNs:", df.shape)

print("\nClass distribution (0 = legit, 1 = fraud):")

print(df["Class"].value_counts())

if "Time" in df.columns:

    df = df.drop(columns=["Time"])

scaler = StandardScaler()

df["Amount"] = scaler.fit_transform(df[["Amount"]])

X = df.drop(columns=["Class"])

y = df["Class"]

X_train, X_test, y_train, y_test = train_test_split(

    X, y,

    test_size=0.3,

    random_state=42,

    stratify=y

)

print("\nTrain class distribution:")

print(y_train.value_counts())

print("\nTest class distribution:")

print(y_test.value_counts())

print("\nBefore SMOTE on y_train:")

```

```

print(y_train.value_counts())

smote = SMOTE(random_state=42, k_neighbors=1)

X_train_sm, y_train_sm = smote.fit_resample(X_train, y_train)

print("\nAfter SMOTE on y_train:")

print(y_train_sm.value_counts())

models = {

    "Logistic Regression": LogisticRegression(max_iter=1000, n_jobs=-1),

    "Random Forest": RandomForestClassifier(

        n_estimators=100,

        random_state=42,

        n_jobs=-1

    ),

    "ANN (MLPClassifier)": MLPClassifier(

        hidden_layer_sizes=(32, 16),

        activation="relu",

        max_iter=50,

        random_state=42

    )

}

results = {}

roc_curves = {}

for name, model in models.items():

    print(f"\nTraining {name}...")

```

```

model.fit(X_train_sm, y_train_sm)

y_pred = model.predict(X_test)

if hasattr(model, "predict_proba"):
    y_prob = model.predict_proba(X_test)[:, 1]
else:
    y_prob = y_pred

acc = accuracy_score(y_test, y_pred)

prec = precision_score(y_test, y_pred, zero_division=0)

rec = recall_score(y_test, y_pred, zero_division=0)

f1 = f1_score(y_test, y_pred, zero_division=0)

try:
    auc = roc_auc_score(y_test, y_prob)
except ValueError:
    auc = np.nan

results[name] = {
    "Accuracy": acc,
    "Precision": prec,
    "Recall": rec,
    "F1-score": f1,
    "ROC-AUC": auc
}

if len(np.unique(y_test)) == 2:
    fpr, tpr, _ = roc_curve(y_test, y_prob)

```

```

roc_curves[name] = (fpr, tpr)

print("\n=== Performance Summary ===")

for name, metrics in results.items():

    print(f"\n{name}:")

    for m_name, value in metrics.items():

        print(f"  {m_name}: {value:.4f}")

plt.figure(figsize=(10, 4))

plt.subplot(1, 2, 1)

y_train.value_counts().plot(kind="bar")

plt.title("Class Distribution (Before SMOTE)")

plt.xticks(rotation=0)

plt.subplot(1, 2, 2)

y_train_sm.value_counts().plot(kind="bar")

plt.title("Class Distribution (After SMOTE)")

plt.xticks(rotation=0)

plt.tight_layout()

plt.show()

if roc_curves:

    plt.figure(figsize=(7, 6))

    for name, (fpr, tpr) in roc_curves.items():

        plt.plot(fpr, tpr, label=f"{name} (AUC={results[name]['ROC-AUC']:.3f})")

    plt.plot([0, 1], [0, 1], "--")

    plt.xlabel("False Positive Rate")

```

```
plt.ylabel("True Positive Rate")  
plt.title("ROC Curves - Fraud Detection")  
plt.legend()  
plt.grid(True)  
plt.show()  
else:  
  
    print("\nROC curves not available (test set may contain only one class).")
```