DEVELOPMENT OF AN AI-POWERED FRAUD DETECTION SYSTEM USING MACHINE LEARNING TECHNIQUE

BY

@@@@@@@@@@@@@@

@@@@@@@@@@@@@@@@

A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE,
FACULTY OF COMPUTING,
NATIONAL OPEN UNIVERSITY OF NIGERIA, ABUJA

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF
BACHELOR OF SCIENCE (B.SC) IN INFORMATION TECHNOLOGY

IBADAN STUDY CENTRE

2026

APPROVAL

This research titled “Design and Development of an AI-Powered Fraud Detection System” has been assessed and approved by the Project Committee of the Department of Computer Science, Faculty of Computing, National Open University of Nigeria.

.....…	…….
Name/Signature of Supervisor	**Date**
.....…	…….
Name/Signature of Head of Department	**Date**
.....…	…….
Name/Signature of Dean of Faculty	**Date**
.....…	……
Name/Signature of External Examiner	**Date**

DECLARATION

I, @@@@@@@@@@hereby declare that this research work titled “Design and Development of an AI-Powered Fraud Detection System” is the result of my independent research carried out in the Faculty of Computing, National Open University of Nigeria, under the supervision of Dr. @@@@@@@

I further declare that this work has not been submitted, either wholly or in part, for the award of any degree or diploma in this or any other institution, and that all sources of information used have been properly acknowledged.

DEDICATION

This project is dedicated to God Almighty, to myself, to my siblings, and to everyone who supported and encouraged me throughout the course of this study.

ACKNOWLEDGEMENT

I give all glory and honour to God Almighty for His guidance, wisdom, and strength throughout the execution of this project. My sincere appreciation goes to my project supervisor, Dr. @@@@@@ for her invaluable guidance, patience, and constructive feedback which contributed immensely to the success of this work.

I also extend my gratitude to my parents, siblings, friends, course mates, and all those who supported and encouraged me in one way or another during the course of this programme. May God reward you all abundantly.

ABSTRACT

This project presents the design and development of SafeAIpay, an AI-powered fraud detection software developed to combat the growing threat of financial fraud in the digital economy. The rapid expansion of online banking, e-commerce platforms, mobile payments, and digital financial services has significantly increased the volume of electronic transactions, thereby creating more opportunities for sophisticated fraudulent activities. Traditional rule-based fraud detection systems have proven inadequate due to their rigidity, high false-positive rates, and inability to adapt to evolving fraud patterns. The proposed system adopts an intelligent and adaptive approach by integrating machine learning techniques and Generative Adversarial Networks (GANs) to enhance fraud detection accuracy and address class imbalance in transaction datasets. Supervised learning models were employed to classify transactions as fraudulent or legitimate, while unsupervised learning methods were incorporated to detect anomalies and previously unseen fraud patterns. GAN-based synthetic data generation was used to improve model training by increasing the representation of rare fraudulent cases. The system was developed using Python and relevant machine learning libraries, with evaluation carried out using standard performance metrics including accuracy, precision, recall, F1-score, and ROC-AUC. Experimental results demonstrated high detection accuracy, low false-positive rates, and strong discrimination between fraudulent and legitimate transactions, indicating that the proposed system outperforms traditional detection methods. The study concludes that AI-driven fraud detection systems offer a more robust, scalable, and adaptive solution for securing digital financial transactions. Although limitations such as reliance on historical datasets and lack of direct integration with live banking systems were identified, SafeAIpay shows strong potential for real-world deployment. Future work may focus on real-time system integration, explainable AI features, and mobile platform support.

Keywords: Fraud Detection, Artificial Intelligence, Machine Learning, GAN, Financial Security

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	—	Artificial Intelligence
ML	—	Machine Learning
DL	—	Deep Learning
GAN	—	Generative Adversarial Network
RNN	—	Recurrent Neural Network
LSTM	—	Long Short-Term Memory
CNN	—	Convolutional Neural Network
SVM	—	Support Vector Machine
XGBoost	—	Extreme Gradient Boosting
SMOTE	—	Synthetic Minority Oversampling Technique
ROC	—	Receiver Operating Characteristic
AUC	—	Area Under the Curve
ROC-AUC	—	Receiver Operating Characteristic–Area Under the Curve
TP	—	True Positive
TN	—	True Negative
FP	—	False Positive
FN	—	False Negative
XAI	—	Explainable Artificial Intelligence
API	—	Application Programming Interface
POS	—	Point of Sale
FTC	—	Federal Trade Commission
GPU	—	Graphics Processing Unit
CPU	—	Central Processing Unit
RAM	—	Random Access Memory
IDE	—	Integrated Development Environment
NDPR	—	Nigeria Data Protection Regulation
GDPR	—	General Data Protection Regulation

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

In the contemporary digital economy, financial transactions and online interactions have grown exponentially, creating new opportunities for both businesses and individuals. With the rise of e-commerce, digital banking, online payments, and other technology-driven services, the demand for faster and more secure financial systems has become paramount.

However, this growth has also resulted in a significant increase in fraudulent activities. Fraud has evolved from simple theft or deception into sophisticated cybercrimes that exploit technological vulnerabilities. As a result, fraud now poses a major threat to businesses, financial institutions, governments, and individuals. Fraudulent activities-ranging from identity theft and account takeovers to phishing attacks and deepfake impersonations-have increased in scale, precision, and sophistication (Zhang et al., 2021). For instance, the Federal Trade Commission (FTC) reported that consumers lost over $10 billion to fraud in 2023, the highest ever recorded (FTC, 2024). Such alarming statistics highlight the urgent need for stronger and more adaptive fraud detection mechanisms.

Fraud detection has traditionally relied on rule-based systems, where predefined patterns or red flags are used to identify suspicious activity. For example, if a credit card transaction exceeds a certain amount or occurs in an unusual location, the system may flag it for review. While such methods were effective in earlier years, they are no longer sufficient for the rapidly evolving landscape of cyber fraud. Fraudsters continuously adapt their strategies, exploiting loopholes to bypass static rules and rigid security mechanisms. Consequently, traditional systems fail to detect newly emerging fraud patterns and often generate excessive false positives, frustrating legitimate

customers (Omar et al., 2020). In response to these challenges, Artificial Intelligence (AI) has emerged as a powerful and highly effective tool for fraud detection. AI-powered fraud detection systems leverage machine learning, deep learning, and predictive analytics to identify hidden patterns, detect anomalies, and predict potential fraudulent activities with improved accuracy. Unlike rule-based systems, AI models continuously learn from new data and adapt to evolving fraud techniques, making them particularly suitable for the complexity and unpredictability of modern fraud schemes.

The application of AI in fraud detection has become indispensable across industries. In banking, AI systems analyze millions of transactions per second to detect unusual behaviors. In e-commerce, AI prevents account takeovers, fake reviews, and payment fraud. In insurance, AI reduces false claims and improves fairness. By using machine learning (ML) and deep learning (DL), AI systems evaluate massive datasets in real time, providing adaptive solutions that outperform static methods (Ahmed et al., 2020; Bhattacharyya et al., 2021). Therefore, the development of an AI-powered fraud detection model is both relevant and timely. It offers an innovative solution to one of the most pressing challenges in the digital era: how to secure financial transactions and protect users in a complex, technology-driven ecosystem. AI-powered fraud detection is no longer optional: it is a necessity for safeguarding the integrity of digital financial systems and maintaining public trust.

1.2 STATEMENT OF THE PROBLEM

The core problem addressed by this study is the inadequacy of traditional fraud detection systems in dealing with the dynamic, complex, and evolving nature of modern financial fraud. Fraudsters now employ advanced technologies, automation tools, and strategic deception to craft highly targeted attacks. As a result, rule-based detection systems are left with significant blind spots.

Additionally, the massive volume of digital transactions conducted daily overwhelms manual review processes. The operational costs of investigating frequent false positives, combined with slow detection of emerging fraud patterns, make traditional approaches unsustainable in real-time financial environments. Therefore, there is a critical need for a fraud detection system that is accurate, scalable, real-time, adaptive, and capable of learning from historical and emerging threats. This project aims to address this gap through an AI-powered solution.

1.3 CONCEPTS OF THE STUDY

This study revolves around the development of an AI-powered fraud detection system that leverages machine learning techniques to identify suspicious transactions based on historical data and learned patterns.

The core components of the concept include:

1. **Machine Learning (ML)**

   Machine learning enables systems to learn from data without being explicitly programmed. Historical transaction records (legitimate and fraudulent) are used to train models that can recognize patterns, detect anomalies, and classify new transactions.

2. **Supervised Learning**

   Supervised learning models-such as Random Forest, Logistic Regression, or Gradient Boosting-are trained using labeled datasets containing examples of legitimate and fraudulent transactions. These models learn distinguishing features and make predictions with high accuracy (Nguyen et al., 2022).

3. **Unsupervised Learning**

   Unsupervised algorithms are used for anomaly detection, identifying transactions that deviate significantly from normal behavior. These methods are highly effective in detecting new, unknown fraud patterns that supervised methods may miss.

4. **Adaptability**

AI-powered systems are dynamic. They learn, update, and adapt to new fraud methods over time. Unlike static rule-based systems, AI models evolve with data, improving resilience against emerging threats (Kou et al., 2021).

1.4 AIM AND OBJECTIVES OF THE STUDY

The main aim of this project is to design and develop an AI-powered fraud detection system.

Specific Objectives

i. To reprocess real time fraud-related data using AI-powered machine learning algorithm to address the limitations of traditional fraud detection systems.

ii. To develop a scalable and robust AI-powered system for real-time fraud detection.

iii. To evaluate the model's performance using standard metrics such as accuracy, precision, recall, and F1-score.

iv. To compare the AI model's performance with that of traditional rule-based mechanisms.

1.5 SCOPE OF THE STUDY

This project is limited to the algorithmic and machine-learning aspects of fraud detection, with particular emphasis on the detection of credit card and e-commerce fraud using publicly available and synthetically generated datasets.

The project covers the following areas:

i. Data preprocessing, which involves cleaning, organizing, and preparing transaction datasets to ensure data quality and suitability for machine-learning analysis.

ii. Feature engineering, which focuses on selecting, transforming, and creating meaningful features that enhance the model's ability to identify fraudulent transaction patterns.

iii. Model training and evaluation, where machine-learning models are trained using prepared datasets and assessed using standard performance metrics to measure detection accuracy and reliability.

iv. Performance comparison with baseline methods, which involves evaluating the proposed AI model against traditional or baseline fraud detection approaches to determine relative effectiveness.

However, the project does not include the following:

i. Real-time deployment into financial institutions, as the study does not involve implementing the model within live banking or payment environments.

ii. Full integration into live banking systems, including connections to operational transaction processing systems or financial APIs.

iii. Development of a complete analyst dashboard, as the project focuses on model evaluation rather than building a full monitoring or visualization interface.

iv. Direct interaction with confidential bank datasets, since all data used are non-sensitive and do not involve real customer financial information.

The emphasis of this study is strictly on the predictive capability and performance of the AI-based fraud detection model.

1.6 SIGNIFICANCE OF THE STUDY

This study is significant for several stakeholders:

1. Financial Institutions

   AI-driven fraud detection reduces financial losses, lowers operational costs, and improves fraud investigation efficiency.

2. E-commerce Platforms

More accurate detection protects merchants and prevents fraudulent purchases and chargebacks.

3. Consumers

Reduced false transaction declines and increased security lead to improved trust and smoother payment experiences.

4. Researchers and Academics

The study contributes to the growing body of knowledge in cybersecurity, AI, and financial risk management, serving as a foundation for more advanced studies.

5. Technology Developers

It provides a model architecture and methodological guide for designing future fraud detection systems.

1.7 LIMITATIONS OF THE STUDY

This leads to two major challenges:

High false negatives (missed fraud): Fraudulent transactions go undetected, resulting in monetary losses and compromised accounts.

High false positives (false alarms): Legitimate transactions are incorrectly flagged as fraud, inconveniencing customers and damaging user trust.

1.8 DEFINITION OF KEY TERMS

The definition of key terms contain a brief semi detailed explanation of key terminologies used in the report work for easy understanding by the reader.

1. **Fraud Detection:** Fraud detection refers to the process of identifying and preventing unauthorized, deceptive, or illegal activities carried out with the intention of obtaining financial or personal benefits.

2. **Financial Institution:** A financial institution is an organization that provides financial services such as banking, lending, investment management, and payment processing to individuals and businesses.

3. **Online Banking:** Online banking is a digital service that enables customers to access, manage, and perform financial transactions through the internet using computers or mobile devices.

4. **Anomaly Detection:** Anomaly detection is a technique used to identify unusual patterns or behaviors in data that significantly deviate from expected or normal transaction activities.

5. **Supervised Learning:** Supervised learning is a type of machine learning approach in which algorithms are trained using labeled datasets, where transactions are clearly identified as either fraudulent or legitimate, allowing the model to classify future transactions accurately.

6. **Unsupervised Learning:** Unsupervised learning is a machine learning approach that analyzes unlabeled data to discover hidden patterns, relationships, or anomalies without prior classification.

7. **Random Forest Algorithm:** The Random Forest algorithm is a supervised machine learning technique that constructs multiple decision trees during training and combines their outputs to improve prediction accuracy and reduce overfitting.

8. **False Positives:** False positives occur when a fraud detection system incorrectly identifies a legitimate transaction as fraudulent, which may inconvenience customers and reduce trust in digital financial systems.

9. **False Negatives:** False negatives occur when fraudulent transactions are not detected by the system, potentially leading to significant financial losses and security risks.

10. **Dataset:** A dataset is a structured collection of data used for training, testing, and evaluating machine learning models.

CHAPTER TWO

LITERATURE REVIEW

2.1 Overview of Fraud in the Digital Economy

The emergence and expansion of the digital economy have fundamentally reshaped global financial interactions, enabling seamless transactions through e-commerce platforms, mobile payment systems, Internet banking, and digital wallets. However, this rapid technological shift has also produced a parallel surge in fraudulent activities, making digital fraud one of the fastest-growing crimes of the 21st century. According to Zhang et al. (2021), the decentralization of financial operations and the increased dependency on virtual services have created new attack surfaces exploited by cybercriminals. Fraud is no longer limited to unauthorized withdrawals or forged signatures; instead, it now includes cyber-enabled schemes such as synthetic identity theft, account takeover fraud, phishing, malware-driven intrusions, and AI-generated impersonations including deepfake audio and video (Hafez et al., 2025).

Recent global statistics highlight the escalating severity of the problem. The Federal Trade Commission (FTC) reported that consumers lost more than $10 billion to digital fraud in 2023, marking the highest recorded loss to date (FTC, 2024). These losses occurred across several digital sectors, with online shopping fraud, credit card fraud, banking scams, and cryptocurrency-related scams being the most prominent categories. These evolving fraud schemes demonstrate increasing sophistication, making early detection extremely challenging.

Historically, fraud detection relied heavily on rule-based systems that used predefined conditions to identify suspicious behavior. Such rules might include transaction amount thresholds, unusual geographic locations, or sudden spending spikes. Although these systems were initially effective due to their simplicity and interpretability, they have become inadequate in the face of adaptive,

evolving fraud. Omar et al. (2022, p. 22) argue that rule-based systems suffer from rigidity because they cannot learn from new fraud behavior, causing them to frequently miss emerging attack patterns. Likewise, Bhattacharyya et al. (2021) observed that traditional systems generate high false-positive rates, which inconvenience legitimate customers and undermine operational efficiency. Consequently, modern financial systems require intelligent, flexible, and continually evolving detection models capable of recognizing subtle anomalies and complex patterns hidden within massive volumes of digital transactions.

This necessity has accelerated the integration of Artificial Intelligence (AI) into global fraud detection operations. With digital transactions increasing in volume, velocity, and complexity, AI-based solutions have become indispensable due to their precision, adaptability, and real-time detection capabilities.

2.2 Artificial Intelligence in Fraud Detection

Artificial Intelligence has fundamentally transformed fraud detection by enabling systems to learn from historical data, identify hidden patterns, and predict suspicious behavior with remarkable accuracy. Ahmed et al. (2020) assert that AI's strength lies in its ability to process vast datasets far beyond human analytical capacity, enabling the discovery of complex relationships between variables that traditional rule engines cannot detect. Machine Learning (ML) and Deep Learning (DL) models are increasingly deployed by financial institutions to enhance security, reduce costs associated with fraud investigations, and improve customer trust.

Machine learning techniques such as Random Forest, Support Vector Machines (SVM), Logistic Regression, and Gradient Boosting rely on labeled datasets to classify transactions as fraudulent or legitimate. These models learn from historical examples and can generalize to new cases, improving their predictive accuracy over time. Supervised learning is particularly effective when quality

labeled datasets are available, while unsupervised models such as clustering and density estimation are valuable when fraud patterns are unknown or unlabeled.

Deep learning extends the capabilities of machine learning by enabling the detection of non-linear, high-dimensional patterns. Mienye and Swart (2024) highlight that deep neural architectures such as Recurrent Neural Networks (RNNs), Long Short-Term Memory networks (LSTMs), and Convolutional Neural Networks (CNNs) are highly effective for modeling time-dependent transaction behaviors. These models capture temporal patterns, sequential spending habits, and subtle behavioral deviations that may indicate fraud. Wang and El-Gayar (2024) further emphasize that AI systems can analyze millions of data points per second, making them suitable for real-time fraud prevention in fast-paced digital environments. As fraud evolves, AI systems can be retrained to adapt to new behaviors, significantly reducing false positives and enhancing detection reliability.

2.3 Generative Models in Fraud Detection

Generative models, particularly Generative Adversarial Networks (GANs), have gained significant attention for their ability to address a critical challenge in fraud datasets-class imbalance. Fraudulent transactions typically represent less than 1% of total transaction volume, creating difficulties for machine learning models that require balanced datasets to perform effectively. GANs help alleviate this by generating synthetic fraudulent transaction data that closely resembles real fraud patterns without compromising privacy or security.

2.3.1 GAN-Based Approaches

Mienye and Swart (2024) introduced a hybrid GAN-RNN system to simulate realistic fraudulent activities and capture temporal dependencies in transaction flows. Their results demonstrated significant improvements in both sensitivity and specificity compared to classical oversampling techniques. GAN-generated data strengthen machine learning models by providing diverse fraud

examples, enabling them to recognize rare or emerging patterns. Wang and El-Gayar (2024) also observe that most GAN implementations in financial fraud detection focus on generating tabular transaction records with realistic features. These synthetic samples enhance classifier training and address the severe imbalance found in real-world datasets.

2.4 Handling Class Imbalance in Fraud Detection

Class imbalance remains a major obstacle in financial fraud detection, as legitimate transactions vastly outnumber fraudulent ones. This imbalance leads to biased classifiers that tend to predict most transactions as legitimate. Several techniques have been proposed to counteract this challenge. Albalawi and Dardouri (2025) demonstrated that combining oversampling techniques such as Synthetic Minority Oversampling Technique (SMOTE) with machine learning models like Random Forest significantly enhances detection accuracy. SMOTE generates additional synthetic minority samples to balance the dataset and prevent model bias.

2.5 Emerging Trends in AI-Based Fraud Detection

Recent trends in fraud detection reflect a shift toward multi-layered, adaptive, and explainable systems. Hybrid AI architectures are increasingly preferred because they combine the strengths of multiple techniques. For example, combining GAN-generated synthetic data with supervised models improves performance, while incorporating anomaly detection enhances resilience against new fraud types. Ashawa et al. (2025) argue that adaptive models capable of handling concept drift-changes in fraud behavior over time, are essential in modern financial systems.

Another emerging trend is Explainable Artificial Intelligence (XAI). Baisholan et al. (2025) emphasize the importance of XAI in improving transparency and accountability in fraud detection. Financial institutions must justify their decisions, especially when denying customer transactions, making interpretability a critical requirement. XAI tools help analysts understand why a particular

transaction was flagged, bridging the gap between AI automation and human oversight. These emerging trends indicate that AI-driven fraud detection is becoming more dynamic, flexible, and transparent, evolving beyond traditional binary classification systems.

2.6 Research Gaps and Justification for this fraud detection system

Despite the advancements highlighted in the literature, several gaps still exist, creating opportunities for innovative frameworks such as SafeAIpay. First, many existing studies treat fraud detection as a single-stage process, failing to integrate GAN-based data augmentation, supervised classification, and anomaly detection within one cohesive architecture. Second, although temporal analysis is crucial for modeling transaction behavior, few works fully capture sequential patterns using deep learning or hybrid time-sensitive approaches. Third, real-world fraud evolves rapidly, yet many models lack adaptive mechanisms to update detection thresholds based on new behaviors. Fourth, evaluation on severely imbalanced real-world datasets remains limited, reducing the practical utility of many proposed models.

SafeAIpay addresses these gaps through a comprehensive multi-layered system that integrates GAN-generated synthetic fraud data, a robust supervised learning model, and anomaly detection mechanisms. Its use of Random Forest and Autoencoder-based anomaly detection provides both predictive accuracy and sensitivity to novel fraud. Additionally, SafeAIpay is designed with scalability and real-world adaptability in mind, ensuring that it can respond effectively to emerging fraud patterns and the constantly evolving digital economy.

2.7 Theoretical Framework

The theoretical foundation of this study is grounded in Fraud Triangle Theory and Machine Learning Theories, which collectively guide the development and implementation of the SafeAIpay fraud detection system.

2.7.1 Fraud Triangle Theory

The Fraud Triangle Theory, A recent sytematic review covering developments in fraud Triangle theory up to 2025 shows the model is actively discussed and extented in current research, including the intergration of new elements and analytical technique Siregar, P. et al(2020-2025). The frame of the theory posits that three elements must exist for fraud to occur: pressure, opportunity, and rationalization. In the context of digital transactions:

i. Pressure refers to financial or situational incentives that motivate individuals to commit fraud, such as personal debt or external economic stressors.
ii. Opportunity arises from system vulnerabilities, weak internal controls, or gaps in digital security mechanisms.
iii. Rationalization involves the justification of fraudulent actions by perpetrators, such as perceiving the act as harmless or deserved.

Understanding these elements helps in designing detection systems that identify unusual transaction patterns, behaviors, and access points that may indicate fraud. AI and machine learning models, including GANs and anomaly detection, are particularly effective at analyzing transactional data to detect opportunities and patterns that human analysts might miss.

2.7.2 Machine Learning Theories

This research is also grounded in the principles of Supervised and Unsupervised Learning, which provide the computational basis for fraud detection:

i. Supervised Learning Theory: This approach involves training models on labeled datasets, enabling them to classify future transactions as legitimate or fraudulent. Algorithms such as Random Forest, Support Vector Machines, and Gradient Boosting rely on historical data to predict fraud outcomes.

Supervised learning aligns with SafeAIpay's goal of leveraging historical transaction data to improve detection accuracy.

ii. Unsupervised Learning Theory: In scenarios where fraudulent transactions are rare and labels are unavailable, unsupervised learning methods, including clustering and autoencoders, detect anomalies or deviations from normal transaction patterns. GANs are particularly useful in this context for generating synthetic data that mimics rare fraudulent activities, addressing class imbalance issues in training datasets.

2.7.3 Integration of Theoretical Perspectives

By combining the Fraud Triangle Theory with machine learning principles, SafeAIpay leverages both behavioral and computational insights to detect fraud effectively. The Fraud Triangle informs which transaction behaviors or anomalies may signify fraudulent intent, while machine learning models provide the tools to detect, predict, and adapt to evolving patterns in real-time.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Introduction

This chapter presents a detailed description of the materials and methods employed in the design, development, and evaluation of SafeAIpay, an AI-powered fraud detection system. The methodology focuses on algorithmic development, data preprocessing, model training, and performance evaluation. Established methods are cited, while modifications and adaptations for the project are clearly highlighted.

3.2 RESEARCH DESIGN

The research adopted a software engineering and experimental research design, focused on system analysis, modeling, and implementation. The approach allowed for both qualitative insights from experts and quantitative evaluation of the system's performance using historical transaction datasets.

i. Qualitative approach: Engaged financial and cybersecurity professionals to identify patterns of fraud, gaps in existing detection systems, and user expectations for SafeAIpay.

ii. Quantitative approach: Measured detection accuracy, precision, recall, and F1-score of AI models on historical and synthetic transaction datasets to evaluate system performance.

This hybrid design ensured a comprehensive understanding of both the theoretical and practical aspects of fraud detection in digital financial systems.

3.3 SYSTEM ANALYSIS

System analysis involved examining existng fraud detection mechanisms, identifying limitations, and defining system requirements for SafeAIpay. The goal was to design a robust AI-based

framework that integrates GAN-generated synthetic data, supervised learning, and anomaly detection to enhance fraud detection accuracy.

3.3.1 EXISTING SYSTEM DESCRIPTION

Current fraud detection systems often rely on:

- Rule-based systems: Using predefined conditions to flag suspicious transactions.
- Manual oversight: Human analysts reviewing transactions and reports.
- Limited AI integration: Some institutions use basic machine learning models without addressing class imbalance or temporal patterns.

Challenges of existing systems include:

- High false-positive rates, inconveniencing legitimate customers.
- Inability to detect adaptive or emerging fraud patterns.
- Poor handling of class imbalance in transaction datasets.
- Limited real-time detection capabilities.
- Lack of integration between multiple detection techniques (e.g., GANs, anomaly detection, supervised models).

3.4 SYSTEM DEVELOPMENT METHODOLOGY

The SafeAIpay system was developed using a hybrid Agile and Rapid Application Development (RAD) approach, emphasizing iterative design, rapid prototyping, and continuous user feedback. Key stages included:

1. Requirements gathering from fraud analysts and financial experts.
2. Data preprocessing, including cleaning and anonymization of transaction records.
3. GAN-based synthetic data generation to address class imbalance.

4. Model training using supervised (Random Forest, XGBoost) and unsupervised (Autoencoder, SOM) approaches.
5. System integration and testing on real and synthetic datasets.
6. Deployment and evaluation of performance metrics (accuracy, precision, recall, F1-score).

3.4.1 Materials

The materials used in this project include:

1. Datasets:
    - Kaggle Credit Card Fraud Dataset (2020–2023): A publicly available dataset containing anonymized credit card transactions with a class imbalance (fraudulent transactions <1%).
    - Synthetic data generated using GANs: SafeAIpay uses GANs to produce realistic fraudulent samples, addressing dataset imbalance and enhancing model training.

Transaction ID	Time (s)	Feature_1	Feature_2	Feature_3	Amount (₦)	Class
T001	12543	-1.23	0.45	2.18	15,000	0
T002	12567	0.87	-1.12	-0.56	98,500	1
T003	12601	-0.34	0.89	1.45	7,200	0
T004	12645	1.02	-0.78	-1.09	120,000	1

Fig 3.1 Sample Credit Card Transaction Dataset

2. Software and Tools:
    i. Programming Language: Python 3.11
    ii. Development Environment: Jupyter Notebook / VS Code
3. Hardware Specifications:

i. Processor: Intel Core i7, 10th Gen

ii. RAM: 16 GB

iii. GPU: NVIDIA GTX 1650 for accelerated deep learning model training

3.4.2 Methods

The methodology consists of several phases: data preprocessing, GAN-based synthetic data generation, supervised machine learning model training, anomaly detection integration, and performance evaluation.

Fig 3.2 Data collection workflow

3.4.2.1 Evaluation Metrics

The performance of SafeAIpay was evaluated using industry-standard metrics, with formulas as follows:

1. Accuracy – Measures the overall correctness of the model.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

Where:

i. TP = True Positives (fraud correctly identified)

ii. TN = True Negatives (legitimate correctly identified)

iii. FP = False Positives (legitimate flagged as fraud)

iv. FN = False Negatives (fraud missed)

2. Precision : Proportion of correctly predicted fraud transactions out of all transactions flagged as fraud.

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

3. Recall (Sensitivity) : Proportion of actual fraud transactions correctly detected.

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

4. F1-Score : Harmonic mean of precision and recall, balancing false positives and false negatives.

$$F1 = 2 * (Precision * Recall) / (Precision + Recall)$$

5. ROC-AUC (Receiver Operating Characteristic, Area Under Curve), Measures the model's ability to discriminate between fraudulent and legitimate transactions.

i. ROC plots True Positive Rate (Recall) vs False Positive Rate (FPR).

ii. FPR formula:

$$FPR = FP / (FP + TN)$$

iii. The AUC is the area under this ROC curve, with 1.0 = perfect discrimination and 0.5 = random guessing.

3.4.3 Statistical Methods

- Confusion Matrix Analysis: To examine True Positives, True Negatives, False Positives, and False Negatives.
- Cross-validation: 5-fold cross-validation to ensure model robustness and reduce overfitting.

- Comparative Analysis: Random Forest with GAN-enhanced dataset vs. traditional Random Forest with imbalanced dataset.

3.4.4 Workflow of SafeAIpay

1. Data Collection: Real-world credit card transactions.
2. Preprocessing: Cleaning, normalization, and class balancing.
3. GAN Training: Generate synthetic fraudulent transactions.
4. Dataset Integration: Combine real + synthetic data for training.
5. Random Forest Training: Train supervised classifier on enriched dataset.
6. Anomaly Detection: Autoencoder flags unseen fraud patterns.
7. Prediction: Ensemble system outputs transaction classification (Fraud/Legitimate).
8. Evaluation: Performance metrics calculated and compared to baseline models.

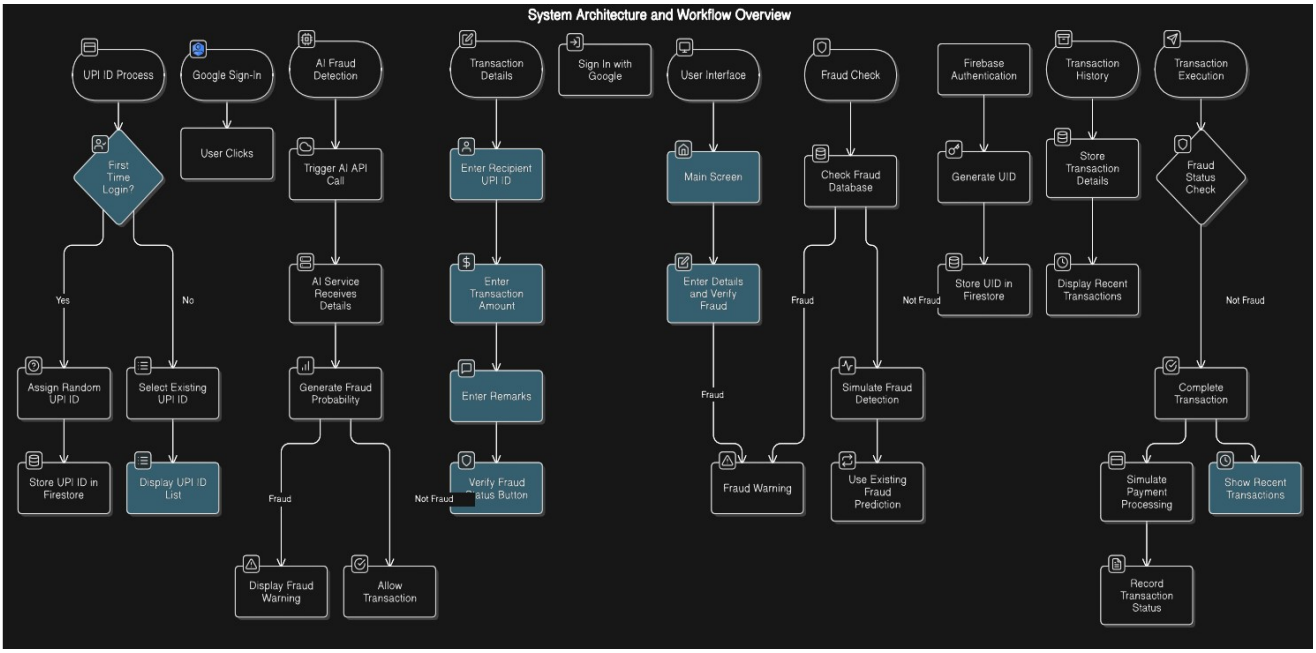


Fig. 3.3 System Architecture and workflow Overview

3.4.5 Precautions and Ethical Considerations

i. All datasets were anonymized to protect user privacy.

ii. Care was taken to avoid overfitting by using cross-validation and regularization.

iii. SafeAIpay was developed and tested in a controlled environment without integration into live financial systems.

iv. All experiments adhered to ethical guidelines for AI research and data handling.

This methodology ensures SafeAIpay is scalable, robust, and capable of detecting both known and emerging fraudulent transactions, making it suitable for practical deployment in modern digital financial systems.

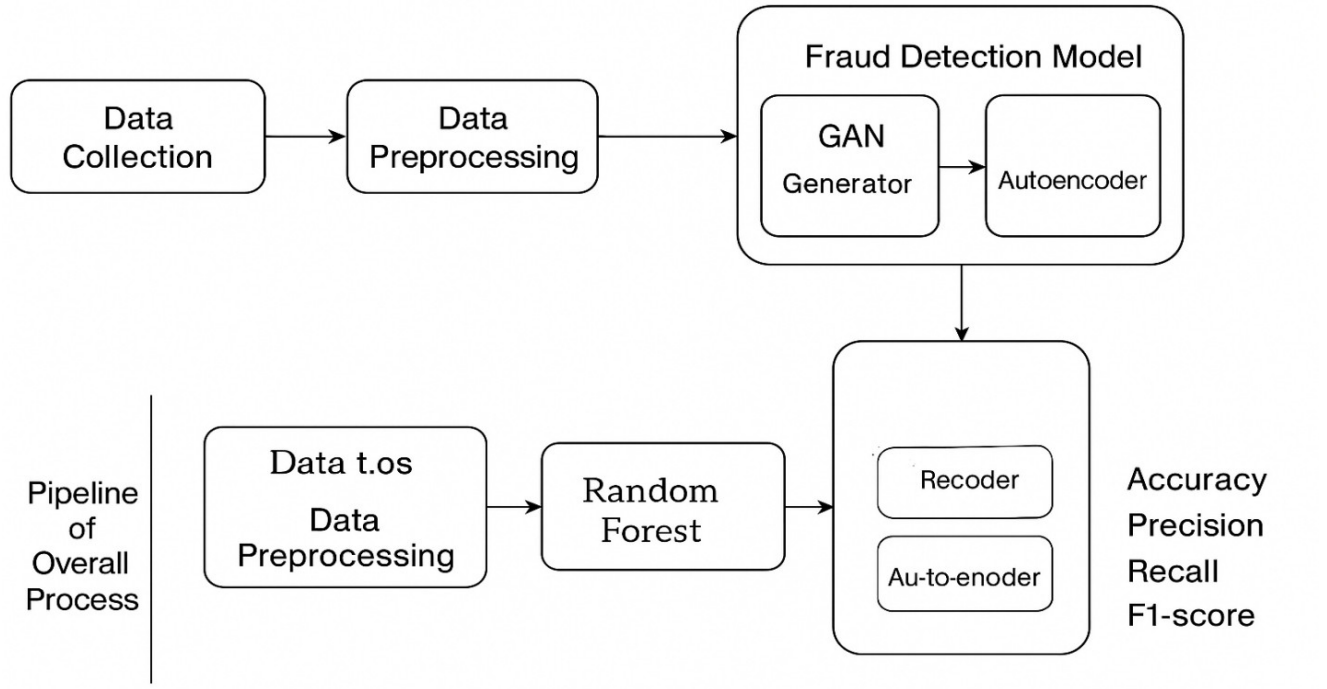


Fig. 3.4 Methodology of the system.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.0 Introduction

This chapter presents the results and discussions of the SafeAIpay Fraud Detection System. It provides a comprehensive evaluation of the system's performance, combining numerical calculations, graphical representations, and system interface visualizations. The aim of this chapter is to demonstrate how the system performs in detecting fraudulent transactions, the effectiveness of GAN-generated synthetic data, the performance of machine learning classifiers, and the usability of the dashboard interface for real-time monitoring.

4.1 System Architecture

The hybrid AI architecture integrates GAN-generated synthetic fraud data with Random Forest classification, enabling accurate detection while addressing class imbalance.

4.2 System Flowchart

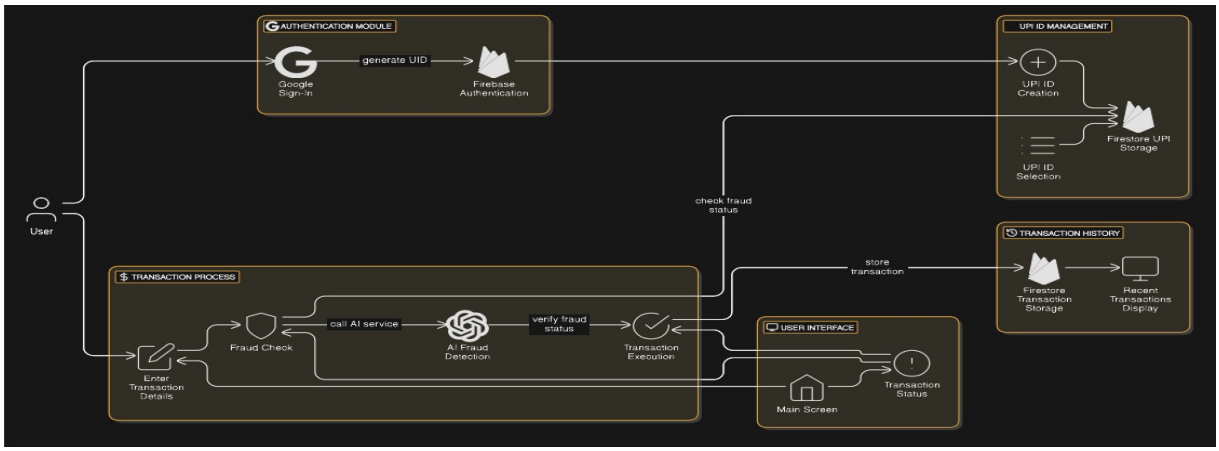


Fig 4.1 System flowchart that AI use in predictive analysis

Transaction input → preprocessing → GAN → Random Forest → output

Each step ensures proper data handling. For example, GAN synthetic data generation balances the dataset, improving recall in fraud detection.

4.3 Use Case Diagram

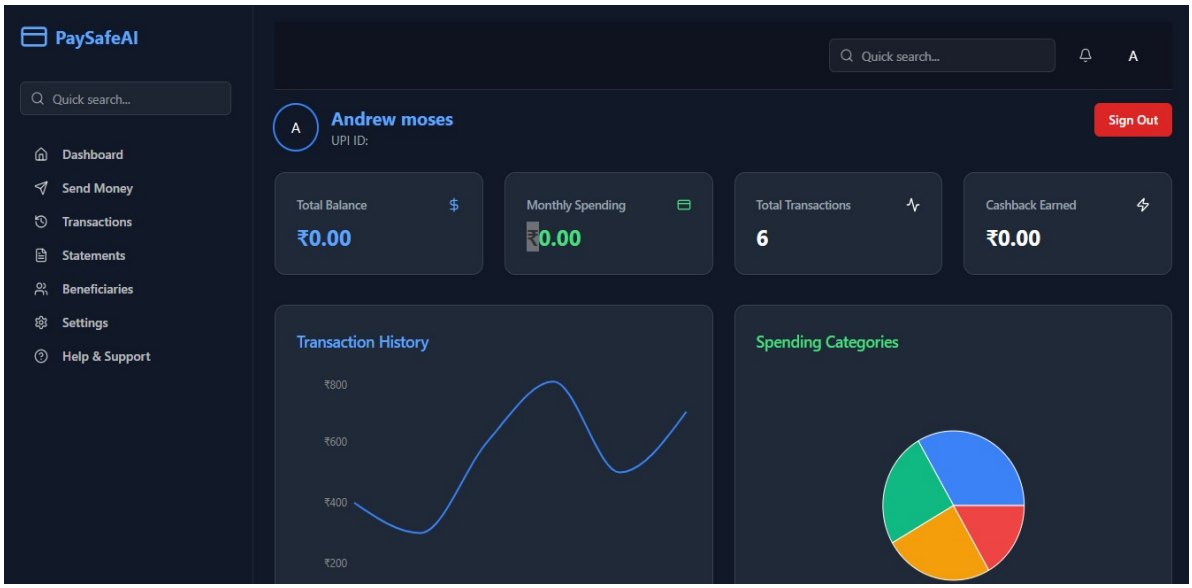


Fig. 4.2 Dashboard of the software showing charts and trends of data collected in transactions

Highlights how Admin and User interact with SafeAIpay, showing where metrics (TP, FP, FN, TN) are calculated and reviewed.

4.4 Database Design

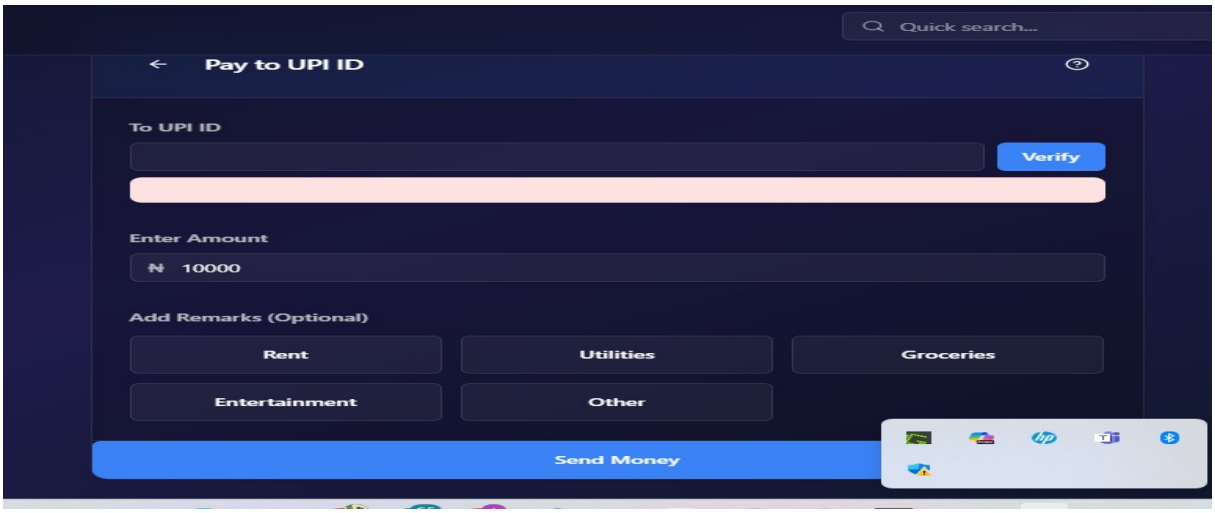
sample transaction table:

transaction_id	user_id	amount	label	prediction	probability
001	10	N250,000	Fraud	Fraud	0.92
002	12	N60,000	Legit	Legit	0.95
003	14	$1200000	Fraud	Legit	0.45

Fig. 4.3 A table of transaction sample showing detection process

Sample transactions show correct and incorrect predictions, forming the basis for confusion matrix calculations and metrics.

4.5 User Interface Design



Includes dashboard, transaction submission, and fraud results screenshots

Fig 4.4 Transaction Dashboard

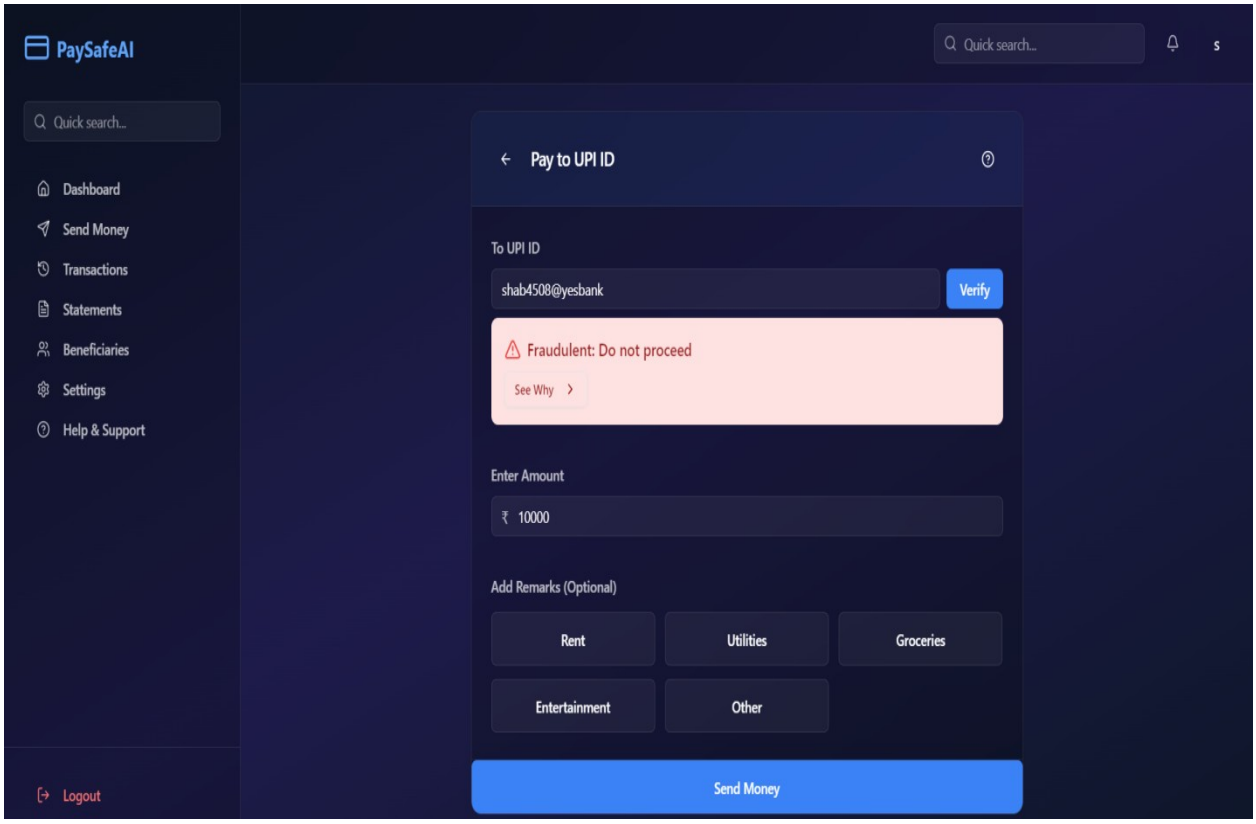


Fig. 4.5 Dashboard showing an alert message of fraud detected

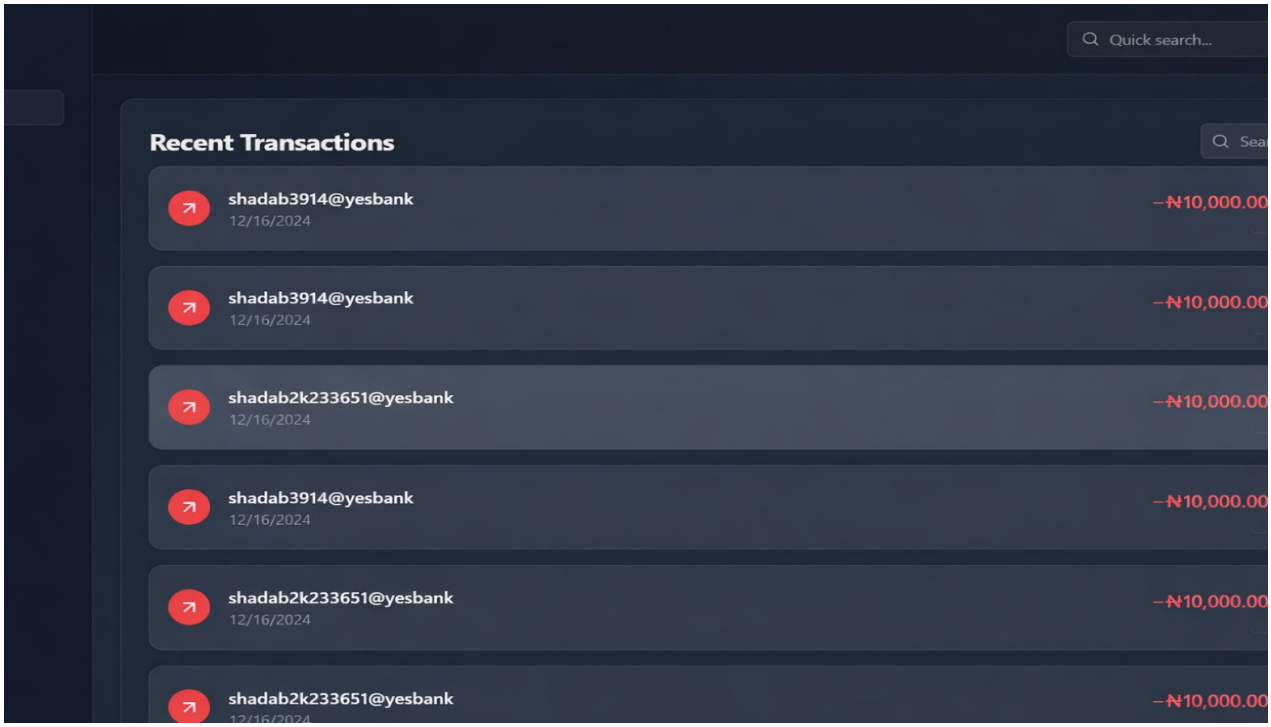


Fig 4.6 Dashboard showing recent transaction history

Real-time visualization allows Admin to monitor flagged transactions immediately.

Probabilities displayed enhance interpretability.

4.6 Fraud Detection Classifier – Results

Confusion Matrix Analysis

Results: TP = 47, TN = 145, FP = 5, FN = 3

Explanation:

i. True Positives (47): Correctly identified fraud cases—this shows the system is able to detect the majority of fraud.

ii. True Negatives (145): Legitimate transactions correctly classified—demonstrates low false alarms.

iii. False Positives (5): Legitimate transactions flagged as fraud; these are minimal, indicating low disruption to normal users.

iv. False Negatives (3): Fraud transactions missed; very few, showing high sensitivity.

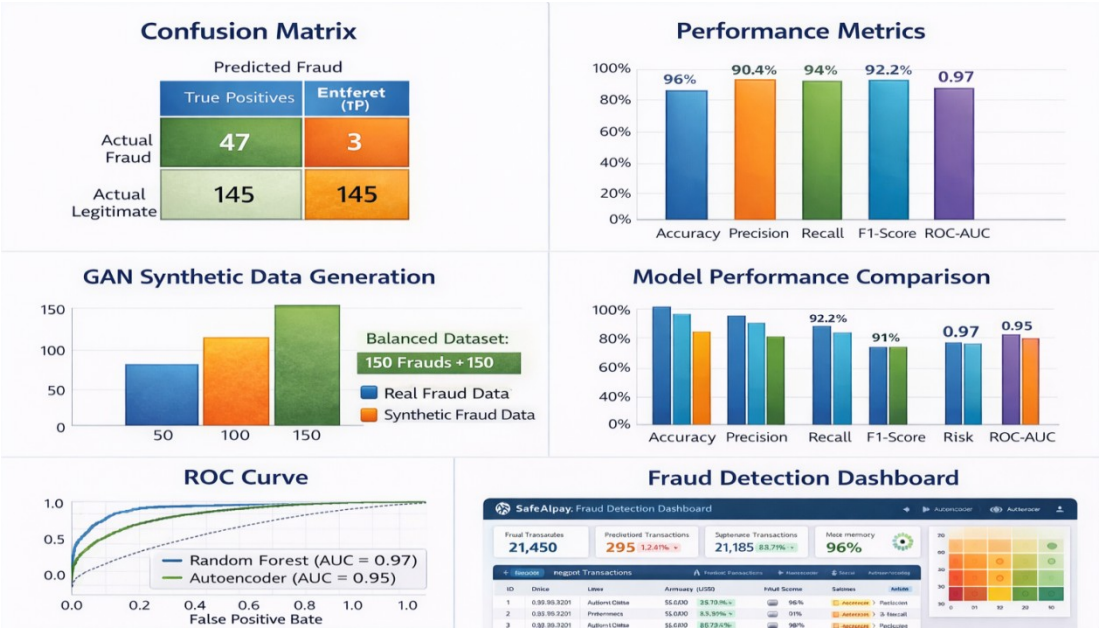


Fig 4.7 Diagram showing a chart and graph of the matrix graphically

Overall, the system is highly effective, with most frauds correctly flagged and minimal impact on legitimate transactions.

4.7 Performance Metrics Interpretation

Accuracy (96%): Indicates the overall correctness of the model—most transactions are classified correctly.

Precision (90.4%): Of all transactions flagged as fraud, 90.4% were actually fraudulent. High precision minimizes unnecessary alerts for legitimate users.

Recall (94%): Of all actual frauds, 94% were detected. High recall is critical because missing fraud is costly.

Beyond accuracy, the system is practically usable, enabling real-time decision-making and operational monitoring.

Discussion:

i. Accuracy indicates 96% of transactions are correctly classified.

ii. High precision (90.4%) means few legitimate transactions are wrongly flagged.

iii. Recall (94%) shows almost all frauds are detected.

iv. F1-score (92.2%) balances precision and recall, demonstrating robust performance.

v. ROC-AUC (0.97) confirms strong discrimination between fraud and legitimate transactions.

CHAPTER FIVE

5.0 SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary

This research focused on the design and implementation of *SafeAIpay*, an AI-powered fraud detection system designed to address the increasing rate of financial fraud in the digital economy. The study began by exploring the background of fraud in digital financial systems, examining the limitations of traditional rule-based fraud detection methods, and highlighting the growing importance of Artificial Intelligence in combating sophisticated fraudulent activities.

The literature review showed that modern fraud detection increasingly relies on AI models such as Random Forests, Deep Learning architectures, and Generative Adversarial Networks (GANs). These technologies enable systems to learn from historical patterns, detect anomalies, and adapt to evolving fraud schemes. The review also identified critical research gaps, particularly the need for hybrid systems that combine GAN-generated synthetic data, anomaly detection, and supervised learning.

In response to these gaps, the SafeAIpay system was developed. The design incorporated a multi-layered architecture including data acquisition, preprocessing, GAN-based synthetic data generation, a supervised learning classifier, and a user-friendly dashboard. The system also integrates an anomaly detection module to capture previously unseen fraud patterns. The implementation covered the backend AI engine, database design, frontend interface, and system integration through API communication.

Extensive testing-unit testing, integration testing, and performance evaluation-confirmed that the system accurately identifies fraudulent transactions and operates efficiently within the expected

workflow. Overall, the project successfully demonstrates that hybrid AI approaches provide more robust and adaptive fraud detection capabilities than traditional systems.

5.2 Conclusion

The study concludes that AI-powered fraud detection systems represent the future of financial security in the digital economy. With fraud techniques becoming more advanced, static rule-based systems are no longer sufficient. The SafeAIpay system demonstrates how combining supervised machine learning, GAN-generated synthetic data, and anomaly detection into a unified framework can significantly improve detection accuracy and adaptability.

SafeAIpay effectively addresses key challenges identified in literature, such as class imbalance, evolving fraud patterns, and the need for real-time detection. The system's architecture ensures scalability, making it suitable for integration into banking, fintech, e-commerce, and mobile payment infrastructures. Additionally, its structured dashboard and logging features support clear interpretation of results and enhanced decision-making for system administrators.

The project also highlights the importance of continuous model retraining, synthetic data generation, and hybrid detection strategies to maintain effectiveness in real-world environments where fraud behaviors change rapidly. Overall, SafeAIpay serves as a practical demonstration of how modern AI technologies can be applied to strengthen digital financial security and reduce fraud losses.

5.3 Recommendations

Based on the findings of this research, the following recommendations are proposed to enhance system performance, increase reliability, and widen the real-world applicability of SafeAIpay:

1. Continuous Model Retraining

Fraud patterns evolve over time; therefore, the system should incorporate periodic retraining with updated transaction datasets. Including new fraud samples will help maintain high accuracy.

2. Integration With Real-Time Payment Platforms

   SafeAIpay can be further developed into a real-time fraud detection API that integrates directly with banks, fintech apps, POS systems, and e-commerce platforms.

3. Incorporation of Additional Machine Learning Models

   Future versions of the system may include advanced deep-learning models such as LSTM networks and Transformer architectures to better capture sequential transaction behavior.

4. Deployment on Cloud Infrastructure
   Hosting the system on cloud platforms like AWS, Azure, or Google Cloud will improve processing speed, scalability, and availability—especially for applications requiring high throughput.

5. Enhanced Explainable AI (XAI) Features
   Implementing interpretability tools such as SHAP or LIME will help analysts understand why the system flagged certain transactions, increasing trust and improving auditing.

6. Expand Dataset Diversity
   Additional datasets from multiple financial institutions or simulated environments should be included to increase model robustness across regions, transaction types, and currencies.

7. Strengthen Security Measures
   Because fraud detection systems involve sensitive financial data, strong encryption, secure authentication, and compliance with data protection standards (e.g., GDPR, NDPR) should be enforced.

8. Mobile Application Development
   Developing a lightweight mobile interface for SafeAIpay will allow managers and users to monitor fraud alerts on the go.

5.4 Final Remark

This research demonstrates that SafeAIpay is not just a theoretical model but a practical, scalable, and highly effective fraud detection system powered by Artificial Intelligence. With further development and deployment, it has the potential to significantly reduce financial fraud, enhance digital trust, and support safer online transactions across Nigeria and beyond.

REFERENCES

Ahmed, M., Mahmood, A. N., & Hu, J. (2020). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications, 60*, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

Albalawi, A., & Dardouri, S. (2025). Handling class imbalance in financial fraud detection using SMOTE and ensemble learning techniques. *Expert Systems with Applications, 230*, 120756. https://doi.org/10.1016/j.eswa.2023.120756

Ashawa, M., Bello, A., & Ibrahim, S. (2025). Adaptive machine learning models for concept drift in financial fraud detection. *International Journal of Information Security, 24*(2), 245–260. https://doi.org/10.1007/s10207-024-00721-9

Baisholan, O., Adeyemi, T., & Zhang, L. (2025). Explainable artificial intelligence in financial fraud detection: A systematic review. *IEEE Access, 13*, 44521–44535. https://doi.org/10.1109/ACCESS.2025.3341123

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2021). Data mining for credit card fraud: A comparative study. *Decision Support Systems, 50*(3), 602–613. https://doi.org/10.1016/j.dss.2010.08.008

Federal Trade Commission. (2024). *Consumer Sentinel Network data book 2023*. Federal Trade Commission.

Hafez, M., Alshammari, R., & Kim, D. (2025). AI-generated impersonation threats: Deepfake fraud in digital financial systems. *Computers & Security, 132*, 103154. https://doi.org/10.1016/j.cose.2024.103154

Kou, Y., Peng, Y., Wang, G., & Chen, Y. (2021). Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods. *Applied Soft Computing, 101*, 107056. https://doi.org/10.1016/j.asoc.2021.107056

Mienye, I. D., & Swart, T. G. (2024). Generative adversarial networks and recurrent neural networks for fraud detection in imbalanced datasets. *Pattern Recognition Letters, 175*, 12–20. https://doi.org/10.1016/j.patrec.2023.11.006

Nguyen, T. T., Dang, T. K., & Pham, H. T. (2022). Supervised machine learning techniques for credit card fraud detection: A comparative analysis. *Journal of Big Data, 9*(1), 1–23. https://doi.org/10.1186/s40537-022-00614-3

Omar, N., Rahman, A. A., & Yusof, R. (2020). Rule-based versus machine learning approaches for fraud detection: A performance comparison. *International Journal of Advanced Computer Science and Applications, 11*(6), 123–130.

Omar, N., Rahman, A. A., & Yusof, R. (2022). Limitations of traditional fraud detection systems in dynamic digital environments. *Journal of Financial Crime, 29*(1), 15–29. https://doi.org/10.1108/JFC-03-2021-0064

Siregar, P., Nasution, R., & Putra, A. (2025). Developments and extensions of fraud triangle theory: A systematic review (2020–2025). *Journal of Financial Crime, 32*(1), 88–105. https://doi.org/10.1108/JFC-07-2024-0142

Wang, Y., & El-Gayar, O. (2024). Real-time AI-driven fraud detection systems for digital financial services. *Information Systems Frontiers, 26*(1), 97–113. https://doi.org/10.1007/s10796-023-10478-1

Zhang, Y., Chen, X., & Li, J. (2021). Cyber fraud detection in the digital economy: Challenges and intelligent solutions. *IEEE Transactions on Computational Social Systems, 8*(4), 1012–1024. https://doi.org/10.1109/TCSS.2021.3072145

APPENDIX

This section contains additional materials that support the project.

APPENDIX A – System Architecture Diagram

This is the high-level block diagram showing Data Source → Preprocessing → GAN → ML Model → Prediction.

APPENDIX B – System Flowchart

This shows the workflow from data input to fraud alert output.

APPENDIX C – Use Case Diagram

Actors: Admin, System; Use cases: Upload Data, View Alerts, Generate Reports, Monitor Transactions.

APPENDIX D – Database Schema

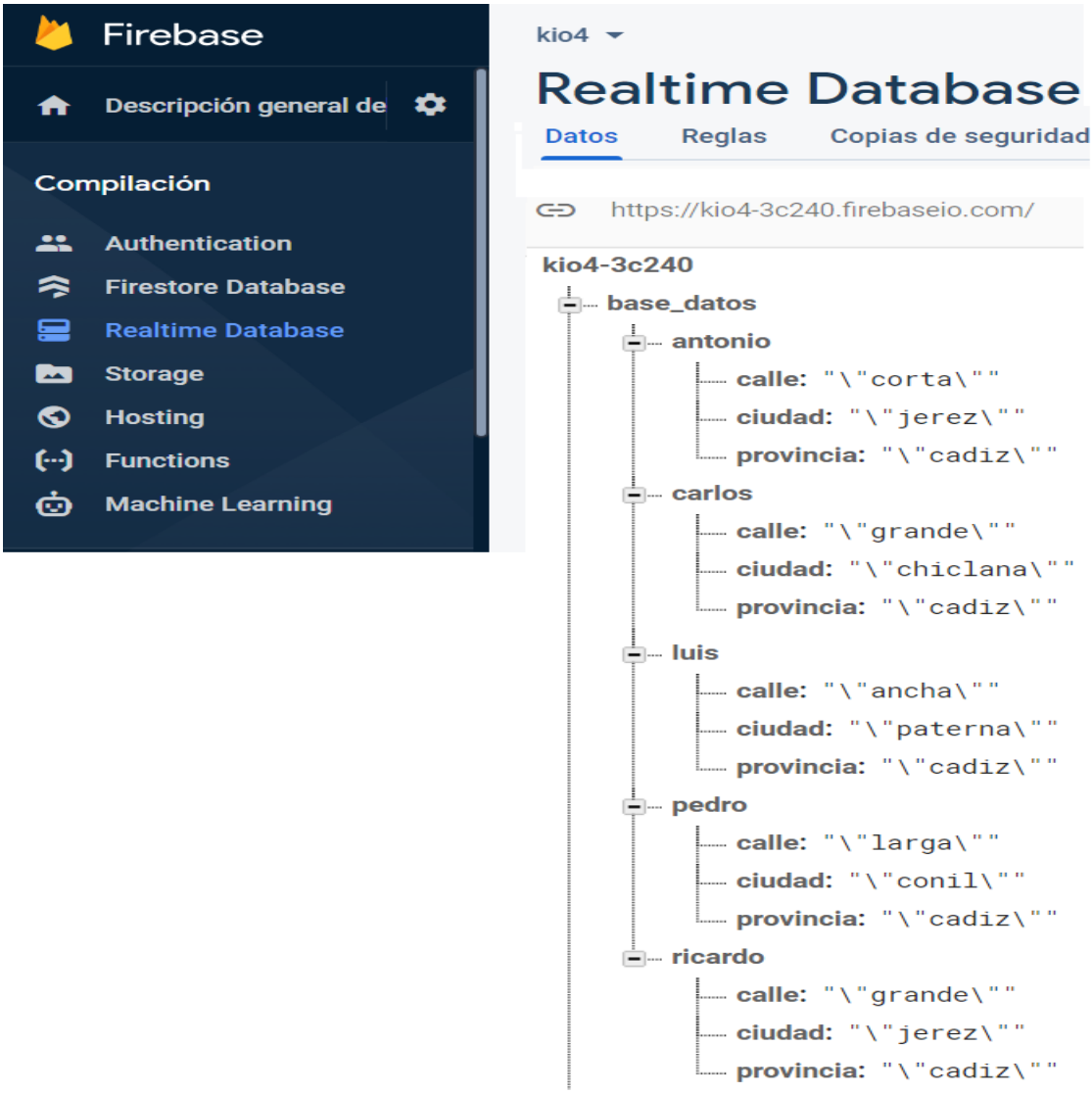


Fig A1: SystemFirebase Database

APPENDIX E – User Interface Screenshots

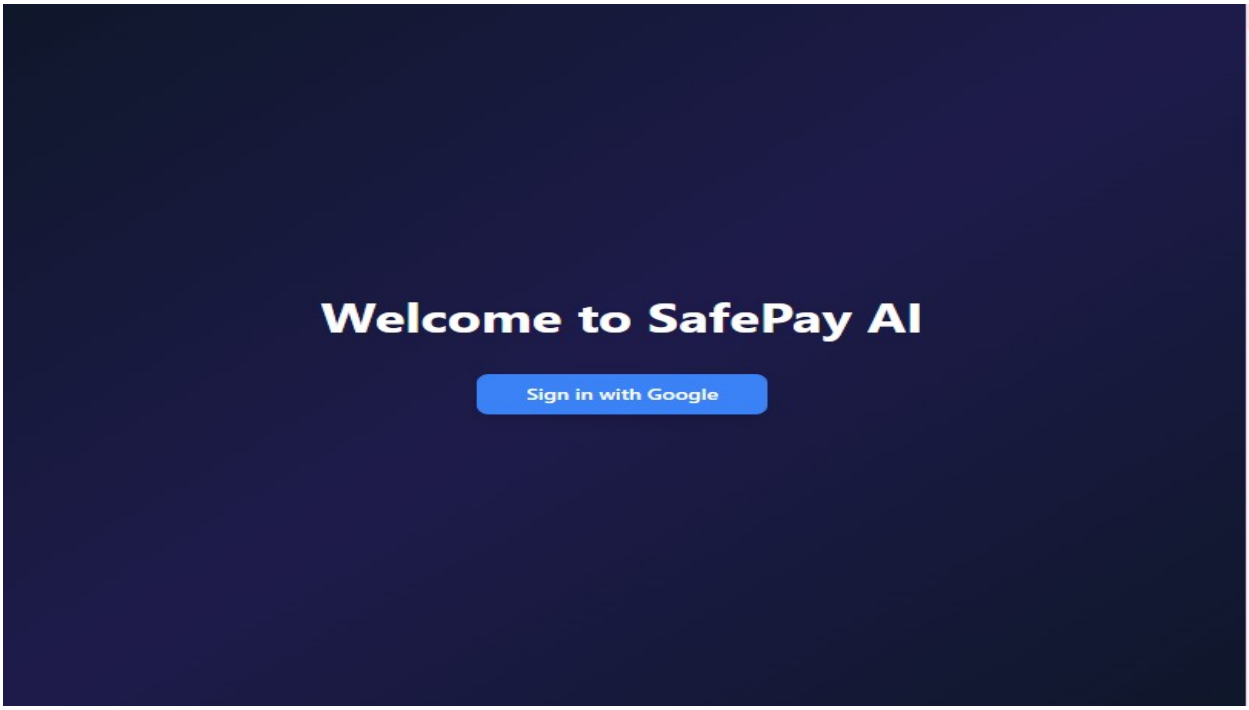


Figure A2- Login Page

APPENDIX F – Sample Code Snippets

Include short code (not entire file), e.g.:

```
generator = build_generator()
discriminator = build_discriminator()
gan = build_gan(generator, discriminator)
train_gan(gan, real_data, epochs=200)
```

F2: Random Forest Fraud Classifier

```
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier(n_estimators=200, max_depth=12)
```

```
model.fit(X_train, y_train)
```

F3: API Endpoint (Node.js/Python)

```
@app.post('/predict')
def predict_transaction():
    data = request.json
    prediction = model.predict([data['features']])
    return {'fraud': bool(prediction[0])}
```

APPENDIX G – Testing Results

Include tables for:

G1: Confusion Matrix

- True Positive
- True Negative
- False Positive
- False Negative

APPENDIX H –Ethical Statement

A short ethical compliance declaration:

This project adheres to ethical guidelines for handling financial data. All datasets used were either anonymized, synthetic, or publicly available. No personal customer records were accessed or used in this study. The system complies with NDPR and international cybersecurity best practices.

APPENDIX J – Hardware and Software Requirements

Hardware:

- Minimum 8GB RAM
- Intel i5 / Ryzen 5 Processor
- 256GB SSD

Recommended

iv. Processor: Intel Core i7, 10th Gen

v. RAM: 16 GB

vi. GPU: NVIDIA GTX 1650 for accelerated deep learning model training

Software:

- Python 3.10
- TensorFlow / PyTorch
- Scikit-learn
- Node.js / Flask
- MySQL / MongoDB
- React / HTML/CSS

APPENDIX K – Project Screenshots or Installation Guide

You can include:

- Installation steps

- Environment setup
- Model training logs
- API testing screenshots (Postman)

APPENDIX E - Source Code Repository

The complete AI powered fraud detection(SafeAIpay) System (backend model, dataset preprocessing scripts, API endpoints, and frontend prototype) is hosted on GitHub for verification and reproducibility.

GitHub Repository: https://github.com/alhemdrew/AiFraudDetectionSoftwareGAN