School of Computing, Engineering and Digital Technologies

Department of Computing and Games

Teesside University

Middlesbrough TS1 3BA

Web-based Credit Card Fraud Detection system using Machine Learning Model

Submitted in partial requirements for the degree of MSc Computer Science (with advanced practice)

May, 2025

Obisanya, Idowu Isaiah

# **ABSTRACT**

There has been a rapid shift in how business operates. Due to the shift to online business such as ecommerce there is a growth in the use of online payment for purchase of goods and services such as purchase of groceries, payment for utility services (such as TV, internet subscriptions etc). Hence, there is a growth in adoption of credit card which then, leads to an escalation of fraud.

Technology has proved to be very effective in different areas of human endeavours. Organizations are adopting the use of machine learning to combat the incident of fraud in credit card transactions.

This project report explores five (5) machine learning models for detecting fraudulent credit-card transactions with their accuracy scores i.e.: Neural Network model (**99.05%);** K-Nearest Neighbour classifier (**99.97%)**; Logistic Regression Classifier (**94.21%**); XGBoost model (**99.91%**) and Random Forest classifier (**99.99%**). Memory utilization, execution time, among other evaluation criteria of each model were observed and noted. The best machine learning model; Random Forest classifier was adopted based on various evaluation criteria and system performance. The adopted model was then deployed through a web-based system where user can input details of a credit card transaction. The system promptly flags to the user if the transaction is fraudulent or valid credit card transaction with high degree of accuracy.

In this research, an extensive exploratory data analysis was carried out about the dataset, the imbalanced dataset was corrected through **Synthetic Minority Over Sampling Technique (**SMOTE). Hyper parameter tuning was also employed using GridsearchCV to improve the performance of the machine learning models.

**CHAPTER ONE**

# **1. INTRODUCTION**

Due to the fast adoption of digital payment system, businesses, globally, are faced with processing massive amount of data volume of transactions that are consummated using electronic payment system. In today’s digitization era, the major challenge that organizations are confronted with is ***Fraud.***

In the recent report, total consumer losses to fraud in the year 2024 amounted to $12.5 billion in United States. This is 25% rise when compared to previous year. The most fraud recorded by consumer was investment scams valued $5.7 billion (FTC, 2025).

According to (FTC, 2019), the total complaints on fraud cases reported by consumer was almost three million cases valued approximately $1.48 billion loss to fraud in 2018; a staggering increase of 38% compared to the previous year. Of the various categories of complaints received by FTC, identity theft was among the major complaints and it is ranked as the third most common complaint. Identity theft relating to ***credit card fraud*** on new account increased by 24% when compared to the previous year. As businesses continue to use various channels for transaction payment such as web-payment, mobile banking, magnetic stripe, contactless (through the Near Field Communication technology (NFC) etc. the trend of fraudulent activities relating to the use of credit card too has grown.

Credit card was one of the main products that customers of financial institutions used to enable them carry out daily transactions and payment for essential services such as TV subscription, gas supplies and electricity, shopping in malls etc. This is due to non-availability of funds at that instance. As customers perform transactions using these credit cards, they accumulate points. Before the last quarter of 2005, a sales of about $190.6 billion was generated from credit cards totalling 56.4 million in Canada. (Pooja, et al., 2021)

## **1.1 Research Question**

Can we develop a system that can allow multiple users to detect fraudulent credit card transactions with high degree of accuracy?

## **1.2 Hypothesis**

Implementing a machine learning-based fraud detection system within a web-based platform will greatly enhance user accessibility, promptness and accuracy of the decision-making by the stakeholders compared to a traditional rule-based system.

## **1.3 Objectives of the project**

1. To investigate previous research on machine-learning model

2. To critically evaluate the various machine learning models to establish and choose the most accurate and effective model based on the dataset.

3. To analyze, design and implement a machine-learning model for credit card fraud detection.

4. To recommend the most suitable machine-learning model for deployment to different categories of users

5. To critically evaluate the speed of processing and ease of use of a web-based machine-learning model

6. To explore the possibilities of deploying a machine learning model through a web-based system.

## **1.4 What is Fraud?**

Fraud is defined as any act carried out by someone that relies on deception for a gain. (Association of Certified Fraud Examiners (ACFE), 2025).

According to (OxfordLearnersDictionaries.com, 2025), fraud is defined as the crime of cheating someone to get money or goods illegally. From these definitions, it could be concluded that fraud involves a deliberate act of obtaining money or goods owned by another through deception.

## **1.5 Why do People Commit Fraud?**

In order to curtail the incident of fraud, it would be worthwhile to understand the motive of the perpetrators. The most widely accepted reason for fraud is shown in Fraud Triangle which was proposed by Dr. Donald Cressey (Fig 1).



Figure 1: Fraud Triangle (Association of Certified Fraud Examiners (ACFE), 2025)

Figure 2 Fraud Triangle (ACFE, 2025)

Figure 2 shows the three elements that are usually present before a fraud can take place:

1. **Financial Pressure**: This is when one is faced with pressure such as greed or ease of the financial gain, or when faced with certain negative circumstances such as loss of job, addiction to gambling, huge debts, health challenges etc.
2. **Rationalization**: This is the justification for carrying out the fraud as conceived by the fraudster. The fraudster may justify why the fraud is okay by saying: “I will pay it back later”, “I deserve it”, “No one works as much I do here”, “I am the least paid despite doing so much”, “I am doing it for my family” etc. There may be a feeling of guilt or shame at the long run if the fraud continues unchecked.
3. **Opportunity**: This means when the existing control in the system is weak. For instance, when there is poor oversight or absence of secondary review (otherwise known as check and balance). In such case, the perpetrator usually starts with a small amount of fraud and then, proceed to a higher magnitude.

It is hypothesized that a person is most likely to commit fraud if the ***three*** elements above are present. (Association of Certified Fraud Examiners (ACFE), 2025)