JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

LSTM Model for Credit Card Fraud Detection

[1]Buchi Nzegwu, [2]Nurudeen M. Ibrahim
[1]*Department of Computer Science,*
[2]*Department of Cyber Security,*
Nile University of Nigeria, Abuja, Nigeria

ABSTRACT
The financial system plays a crucial role in managing and transferring money and assets, with credit cards serving as a primary means of transaction. While credit cards offer convenience, they also present security challenges, as fraudsters continuously attempt to exploit vulnerabilities. Detecting fraudulent transactions is essential to mitigating financial losses. This study develops a robust fraud detection model using Long Short-Term Memory (LSTM) networks. While LSTMs have been widely used in credit card fraud detection, our approach introduces key architectural modifications that enhance stability, regularization, and feature transformation. These modifications improve model generalization and performance in detecting fraudulent transactions. Given the highly imbalanced nature of the dataset where fraud accounts for only 0.6% of training data and 0.3% of test data Synthetic Minority Over-Sampling Technique (SMOTE) is employed to generate synthetic samples for the minority class in the training set. The model's performance is evaluated using key metrics such as accuracy, precision, recall, F1-score, and the AUC-ROC curve.

ARTICLE INFO
Article History
Received: March, 2025
Received in revised form: June, 2025
Accepted: July, 2025
Published online: September, 2025

KEYWORDS
Deep learning, Credit Card Fraud, Fraud Detection, Data Imbalance, SMOTE

INTRODUCTION

The financial system remains a fundamental pillar of every economy, facilitating transactions between individuals and businesses. Among the various payment methods, credit cards have played a major role in advancing cashless transactions, making payments more convenient by eliminating the need to carry physical cash. The concept of credit cards was first introduced in February 1950 by Frank McNamara (Swipesum, 2023). Since the emergence of the internet age, credit card usage has gained widespread adoption, particularly in e-commerce, where businesses and consumers can seamlessly interact both online and offline. Customers can easily compare prices, purchase products, and make payments from the comfort of their homes using credit cards.

However, alongside the benefits of this technology, credit card fraud has become a growing concern, particularly with the expansion of digital transactions. Credit card fraud involves the unauthorized use of card credentials, leading to financial losses for individuals, businesses, and financial institutions. According to a report by Statista, global losses due to card fraud including both credit and debit card fraud amounted to approximately 30 billion USD between 2020 and 2022, impacting both merchants and card acquirers (Statista, 2024). Fraudsters employ various techniques to carry out illegal transactions, including phishing for credit card information and counterfeiting cards. Given the increasing sophistication of fraudulent activities, there is a critical need for effective fraud detection mechanisms to safeguard financial transactions.

Traditional fraud detection methods have struggled to combat complex fraud patterns, necessitating more advanced approaches (Cherif et al., 2023). This study aims to develop a robust fraud detection model using Long Short-Term Memory (LSTM) networks to enhance the detection of fraudulent credit card transactions and mitigate financial risks. LSTM networks are an

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

extension of Recurrent Neural Networks (RNNs) designed to overcome the limitations of RNNs, particularly the vanishing and exploding gradient problems. These issues arise during back propagation when gradients either become too small (vanishing) or excessively large (exploding), making it difficult for RNNs to learn long-term dependencies in sequential data. Since credit card fraud datasets often contain time-series data (Staudemeyer & Morris, 2019), LSTM networks are well-suited for modelling fraudulent transaction patterns.

One major challenge in machine learning and deep learning is class imbalance. A dataset is considered imbalanced when one class significantly outweighs another in terms of representation. Class imbalance is common in medical diagnoses, where rare diseases appear far less frequently than healthy cases. Similarly, in credit card fraud detection our focus in this study the number of legitimate transactions is disproportionately higher than fraudulent ones. Researchers classify the problem of class imbalance into multi-class imbalance and binary-class imbalance, with credit card fraud detection falling under the latter category (Thabtah, Hammoud, Kamalov, & Gonsalves, 2020).

Various approaches have been proposed to address class imbalance, including data-level methods, algorithm-level methods, and ensemble methods (Hasib et al., 2020). This study employs a data-level approach, specifically oversampling, to balance the dataset. The technique used is Synthetic Minority Over-Sampling Technique (SMOTE), which generates synthetic samples for the minority class to improve class distribution. SMOTE enhances the model's ability to detect fraudulent transactions by preventing bias toward the majority class.

Credit cards have undoubtedly revolutionized financial transactions, significantly enhancing the efficiency and convenience of payments in the 21st century. However, with every technological advancement come vulnerabilities that malicious actors seek to exploit. Credit card fraud has become a major challenge, posing financial risks to individuals, businesses, and financial institutions. Fraudsters employ sophisticated techniques such as phishing, card skimming, counterfeit card creation, and unauthorized transactions, making it increasingly difficult for traditional fraud detection systems to keep up. Many existing fraud detection models rely on rule-based systems or conventional machine learning algorithms that struggle to detect evolving fraud patterns. Additionally, credit card fraud detection is complicated by the highly imbalanced nature of transaction data, where fraudulent cases constitute only a small fraction of total transactions (Mienye & Jere, 2024). Given these challenges, there is a need for a more robust and intelligent fraud detection system capable of identifying fraudulent transactions with high accuracy while minimizing false positives. This study aims to address this problem by developing a deep learning-based credit card fraud detection model using Long Short-Term Memory (LSTM) networks. The proposed model will leverage time-series transaction data and employ techniques such as Synthetic Minority Over-Sampling Technique (SMOTE) to handle class imbalance, improving the overall effectiveness of fraud detection systems.

Credit card fraud detection has been a widely researched areas in financial technology and cybersecurity due to its significant impact on economic stability. Over the years, various fraud detection techniques have been developed, ranging from traditional rule-based methods to advanced machine learning and deep learning approaches(Mallidi & Zagabathuni, 2021). The review highlights machine learning-based models and deep learning architectures.

Hasib et al. (2020) provides a survey of 24 studies from 2003 to 2019, analysing various approaches to addressing class imbalance in machine learning. It reviews techniques such as data sampling, cost-sensitive learning, Genetic Programming, bagging, and boosting. The study highlights the importance of hybrid and ensemble methods in improving classification performance for imbalanced datasets. Additionally, it presents a statistical analysis of different classification algorithms under diverse conditions. The survey concludes that balancing the dataset is essential for improving classifier performance and

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

emphasizes the role of feature selection in enhancing system efficiency. However, the study lacks an in-depth comparative analysis of the surveyed papers and does not propose a definitive framework for handling class imbalance.

The paper in (Mallidi & Zagabathuni, 2021) evaluates logistic regression, decision trees, and random forest for credit card fraud detection. The results show that random forest outperforms the other models, achieving 96% accuracy and an AUC of 98.9%. Based on these findings, the study recommends random forest as the most effective approach. Additionally, the study identifies individuals above 60 years as the most frequent fraud victims, with fraudulent transactions peaking between 22:00 GMT and 4:00 GMT. These insights highlight the role of behavioural patterns in fraud detection. While the models perform well, they rely on structured features and may struggle with evolving fraud tactics.

Similarly, the study in (Karkaba, Adnani, & Erritali, 2023)investigates the use of artificial neural networks (ANN) and convolutional neural networks (CNN) for fraud detection, acknowledging their efficiency but also highlighting two key challenges: the continuous emergence of new fraud patterns and the presence of highly imbalanced datasets. To address these issues, the study incorporates an autoencoder, an unsupervised deep learning approach, to compare its performance with supervised models.

The study evaluates these models on a dataset of 284,807 credit card transactions, applying various preprocessing techniques such as normalization, data balancing, and feature selection. Additionally, hyperparameter tuning, boosting techniques, and confusion matrix analysis are conducted to optimize model performance. The results show that one model stands out for its ability to detect new fraud patterns, achieving an F1-score of 93%.

The study in (Karkaba et al., 2023) shifts focus to online credit payment fraud detection, emphasizing user behaviour modelling and the structural features of web pages. Traditional methodologies for sequential modelling often fail to capture the underlying structure of web pages, limiting their effectiveness. To address this, the authors propose the SAH-RNN model, which leverages multi-scale behaviour sequences derived from different web page granularities. The model consists of stacked RNN layers where higher layers process aggregated behaviours less frequently while receiving information from lower layers. Additionally, a dual attention mechanism is introduced to capture both sequential dependencies within individual sequences and structural relationships across different levels of web page granularity. Experimental validation using a large-scale real-world transaction dataset from Alibaba demonstrates that SAH-RNN outperforms state-of-the-art approaches in detecting fraudulent transactions. This study underscores the importance of leveraging both user behaviour analysis and web page structures in fraud detection, marking a significant advancement in fraud detection methodologies.

MATERIALS AND METHODOLOGIES

Long Short-Term Memory (LSTM) networks are a specialized type of recurrent neural network (RNN) introduced by Hochreiter & Schmidhuber in 1997, they are designed to address the limitations of traditional RNNs. Standard RNNs are prone to the vanishing and exploding gradient problems, which hinder their ability to retain long-term dependencies. LSTMs overcome this by employing gating mechanisms that regulate the flow of information, enabling the network to capture both short-term and long-term memory. These gates play a critical role in preventing the gradient issues commonly encountered in vanilla RNNs, allowing for more effective learning over extended sequences (Liu & Zhang, 2021).

Recurrent Neural Networks (RNNs) are designed with a chain of repeating modules, allowing them to process sequences by maintaining information about previous events. Unlike standard RNNs, Long Short-Term Memory (LSTM) networks feature a more complex structure incorporating four distinct neural networks within their repeating module chain. This additional complexity enables LSTMs to better

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

capture and retains long-term dependencies in sequences.

Data Pre-processing

The dataset consists of numerical and categorical transaction attributes. To ensure compatibility with machine learning models, categorical variables are encoded using label encoding. Unnecessary attributes, such as personally identifiable information (e.g., customer names and addresses), are removed to prevent data leakage. Timestamp features are transformed to extract relevant components such as year, month, and day, while DOB is used to compute the account holder's age before being discarded. This transformation ensures temporal information is retained while minimizing redundancy.

Synthetic Minority Over-Sampling Technique (Smote) Analysis

The Synthetic Minority Over-Sampling Technique (SMOTE) is an oversampling method that generates synthetic data points to balance the minority class. SMOTE works by generating synthetic samples for the minority class rather than simply duplicating existing instances. The algorithm selects instances from the minority class and generates synthetic data points along the line segments connecting these instances to their nearest neighbours. This approach helps in increasing the minority class by diversification and not merely just duplicating.

Model Development

A Long Short-Term Memory (LSTM) network is employed to capture sequential transaction patterns. The model architecture consists of: Input layer: Corresponding to the number of features in the dataset, Two LSTM layers: Each with 64 hidden units to capture temporal dependencies, Dropout layer (50%): Applied to prevent overfitting, Fully connected output layer: Using a sigmoid activation function for binary classification. The model is trained using Binary Cross-Entropy Loss (BCEWithLogitsLoss) with class-weighted adjustments to mitigate class imbalance. The Adam optimizer is used for weight updates with an initial learning rate of 0.001, and a Step Learning Rate scheduler reduces the learning rate by 0.1 every five epochs to stabilize convergence.

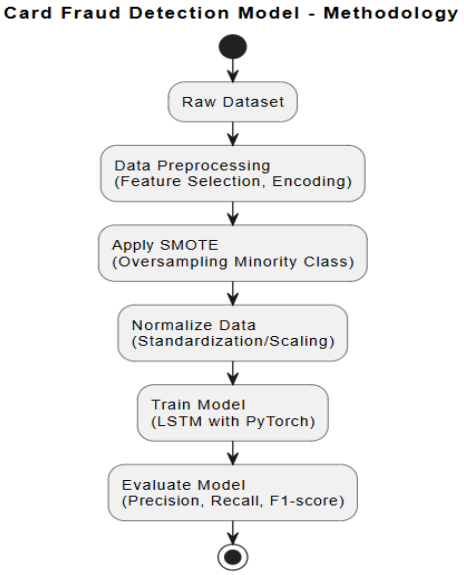

Figure 1: Proposed Credit Card Fraud Detection Model

Dataset

The dataset used in this study was obtained from Kaggle, covering transaction data from January 1, 2019, to December 31, 2020. It is a highly imbalanced dataset, provided in separate CSV files for training and testing. The training set consists of 1,296,675 entries, with 99.5% of the transactions labelled as non-fraudulent and 0.5% as fraudulent. The test dataset contains 555,719 entries, with 99.7% labelled as non-fraudulent and 0.3% as fraudulent.

This study employs a deep learning-based approach for fraud detection in financial transactions. The methodology consists of five key stages: data preprocessing, handling class imbalance, feature scaling, model development, and performance evaluation. Figure- 1 shows the flowchart illustrating the methodology used for fraud detection, including data pre-processing, SMOTE application, normalization, model training, and evaluation.

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

Dataset Analysis

The dataset used for training consists of real-world credit card transaction records, which exhibit a significant class imbalance between fraudulent and non-fraudulent transactions. Table 1 provides a summary of the dataset characteristics before and after pre-processing.

Table 1: Test model specifications and test conditions.

Class	Count (Before Preprocessing)	Count (After SMOTE)
Non-Fraud (Class 0)	1289169	1289169
Fraud (Class 1)	7506	644584

The training data for credit card fraud detection exhibited a high form of imbalance, where non-fraudulent transactions accounted for approximately 99.6% of the dataset, while fraudulent transactions made up only 0.6%. To mitigate this imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied with a sampling strategy of 0.5. This technique generated synthetic data for the minority class, increasing its count to 50% of the majority class. The LSTM model was trained using PyTorch with the following Hyper parameter as shown in Table below

Evaluation Metric

Given the imbalance in fraud detection, traditional accuracy metrics are insufficient. Instead, the model's performance is assessed using the following key metrics:

1. Precision (P): Measures the proportion of predicted fraudulent transactions that are actual fraud.
2. Recall (R): Measures the percentage of actual fraudulent transactions correctly identified.
3. F1-Score: The harmonic mean of precision and recall, providing a balanced assessment of the model's effectiveness.
4. Precision-Recall AUC (PR AUC): Evaluates the trade-off between precision and recall, ensuring the model prioritizes reducing false negatives.

RESULTS

The results from the model indicate that it performs well in addressing the challenges posed by the highly imbalanced nature of fraud detection problems, particularly in the context of credit card fraud. One of the key observations is the successful use of SMOTE (Synthetic Minority Over-sampling Technique) for balancing the training dataset. The primary challenge in fraud detection is the imbalance between the number of fraudulent and non-fraudulent transactions, where the latter vastly outnumbers the former. By oversampling the minority class (fraudulent transactions), SMOTE ensured that the model had more balanced data to learn from, which contributed to the improved recall.

Recall is critical in fraud detection, as it reflects the model's ability to correctly identify fraudulent transactions. In this case, a higher recall ensures that the model is not overlooking fraud cases, which could result in significant financial losses or security breaches. By incorporating SMOTE, we were able to mitigate the risk of the model failing to detect fraud, thereby improving its sensitivity towards fraudulent transactions.

However, the trade-off between precision and recall must also be considered. While high recall ensures fewer fraudulent cases are missed, it may lead to an increase in false positives non-fraudulent transactions that are wrongly classified as fraudulent. This could result in unnecessary transaction flags, increasing the number of legitimate transactions that are incorrectly flagged for further scrutiny. To address this, threshold tuning was performed, which allowed for a careful balance between precision and recall. By adjusting the decision threshold, we were able to optimize the model to minimize false positives without sacrificing its ability to detect fraud.

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

JOURNAL OF SCIENCE TECHNOLOGY AND EDUCATION 13(3), SEPTEMBER, 2025
E-ISSN: 3093-0898, PRINT ISSN: 2277-0011; Journal homepage: www.atbuftejoste.com.ng

The model demonstrated a favourable balance, with high precision and recall, validating its effectiveness as a tool for credit card fraud detection. The precision value indicated that the fraudulent transactions identified by the model are highly likely to be actual frauds, which is crucial for minimizing false alerts.

Table 3: Classification Report

Metric	Non-Fraud (Class 0)	Fraud (Class 1)	Macro Avg	Weighted Avg
Precision	0.9984	0.9540	0.9725	0.9973
Recall	0.9993	0.8635	0.9126	0.9974
F1-Score	0.9989	0.9065	0.9405	0.9973

Performance Visualization

To further analyse the model's predictive capabilities, we present a visualization of its performance metrics. The precision-recall curve (or ROC curve) provides insights into how well the model maintains precision at different classification thresholds. The curve in Figure 2 illustrates the trade-off between recall and precision across varying thresholds, highlighting the model's ability to detect fraudulent transactions effectively.

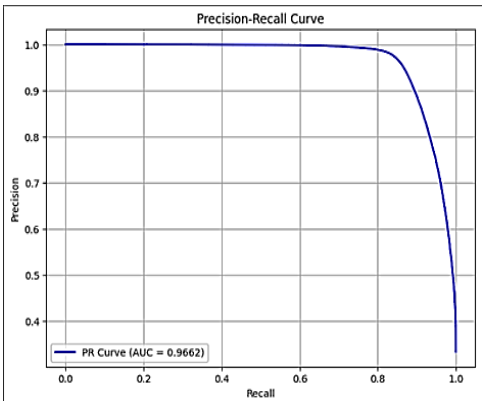

Figure 2: Precision-Recall Curve

CONCLUSION

By incorporating Layer Normalization, ReLU activation, and strategically placed Dropout, our LSTM-based fraud detection model offers a more robust, stable, and generalizable approach compared to conventional architectures. These innovations allow the model to better capture fraudulent transaction patterns while maintaining computational efficiency.

Credit card fraud stills remain a prevalent issue in this era of digitalization and several approaches have and still been adopted to tackle this menace. This model has tried to build on other existing and also serve as a shoulder where upcoming model can stand. From the evaluation LSTM has shown a significant results cumulating the fact it can tackle this menace. The successfully application of SMOTE in this fraud detection model highlights the importance of handling imbalanced datasets when working with real-world problems like credit card fraud detection.

Fraudulent transactions are rare, making it challenging for traditional machine learning models to detect them without incorporating strategies like SMOTE or similar oversampling techniques. Despite the promising results, there is room for further enhancement. Feature engineering could provide additional insights into the characteristics of fraudulent transactions. Features such as transaction time, merchant type, and customer behavior patterns could further improve model accuracy. Additionally, incorporating ensemble models, which combine the predictions of multiple models, may provide better performance and generalizability than a single model. Techniques such as bagging, boosting, or stacking could potentially yield improvements by reducing variance and bias.

REFERENCES

Cherif, Asma, Badhib, Arwa, Ammar, Heyfa, Alshehri, Suhair, Kalkatawi, Manal, & Imine, Abdessamad. (2023). Credit card fraud detection in the era of

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.

disruptive technologies: A systematic review. *Journal of King Saud University-Computer and Information Sciences, 35*(1), 145-174.

Hasib, Khan Md, Iqbal, Md Sadiq, Shah, Faisal Muhammad, Mahmud, Jubayer Al, Popel, Mahmudul Hasan, Showrov, Md Imran Hossain, . . . Rahman, Obaidur. (2020). A survey of methods for managing the classification and solution of data imbalance problem. *arXiv preprint arXiv:2012.11870*.

Karkaba, IMANE, Adnani, EM, & Erritali, M. (2023). Deep Learning Detecting Fraud in Credit Card Transactions. *Journal of Theoretical and Applied Information Technology, 101*(9).

Liu, Kai, & Zhang, Jie. (2021). A dual-layer attention-based LSTM network for fed-batch fermentation process modelling *Computer aided chemical engineering* (Vol. 50, pp. 541-547): Elsevier.

Mallidi, Manoj Kumar Reddy, & Zagabathuni, Yeshwanth. (2021). Analysis of Credit Card Fraud detection using Machine Learning models on balanced and imbalanced datasets. *International Journal of Emerging Trends in Engineering Research, 9*(7).

Mienye, Ibomoiye Domor, & Jere, Nobert. (2024). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. *IEEE Access*.

Staudemeyer, Ralf C, & Morris, Eric Rothstein. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

Thabtah, Fadi, Hammoud, Suhel, Kamalov, Firuz, & Gonsalves, Amanda. (2020). Data imbalance in classification: Experimental evaluation. *Information Sciences, 513*, 429-441.

Swipesum, "The history of the credit card," https://www.swipesum.com/insights/history-of-the-credit-card. Statista, "Global card fraud losses worldwide from 2014 to2023,"https://www.statista.com/statistics/1394119/global-card-fraud-losses/.

Corresponding author: Buchi Nzegwu
nzegwumo@gmail.com
Department of Computer Science, Nile University of Nigeria, Abuja, Nigeria.