A PREDICTIVE MODEL FOR CREDIT CARD FRAUD DETECTION USING DEEP LEARNING

Dr. Venugeetha Y
Professor, Department of Computer Science & Engineering, Don Bosco Institute of Technology Bangalore, India
vgydbit@gmail.com

Anushree B C, Tejaswini S D,Sushmitha M, Kusuma S
Department of Computer Science & Engineering, Don Bosco Institute of Technology Bangalore, India

Abstract
The growth in online transactions has resulted in a major spike in credit card fraud, creating serious challenges for financial institutions across the globe. This study is a detailed analysis of cutting-edge machine learning and deep learning algorithms to improve identification of fraudulent credit card transactions. Different algorithms like Decision Trees, Logistic Regression, Random Forest, Support Vector Machines, k- Nearest Neighbors, Artificial Neural Networks, and Recurrent Neural Networks are investigated and compared on publicly available transaction datasets. Particular focus is placed on addressing class imbalance and model performance optimization using precision, recall, F1-score, and AUC-ROC metrics. Along with data-driven approaches, rule-based validation like Luhn's algorithm and CVV checks are utilized to enhance overall system precision. The research points out that ensemble models and deep learning models provide enhanced accuracy and generalization in detecting fraudulent patterns. The results help improve adaptive, scalable, and convenient fraud detection architectures that can enhance real-time decisions and reduce false positives to foster better security and customer trust in financial systems.

Keywords
Credit card fraud, machine learning, deep learning, fraud detection systems, ANN, RNN, ensemble models, imbalanced datasets

1. Introduction

The growth of online payments and electronic commerce has revolutionized the financial environment, bringing more convenience but also heightening susceptibility to credit card fraud. This form of fraud, encompassing activities like identity theft, phishing, and unauthorized purchases, is a significant threat to individuals and organizations, resulting in large financial losses and decreased confidence in digital payment systems.

Legacy fraud detection systems, based on pre-defined rules and thresholds, are not keeping pace with the dynamic nature of fraud. They tend to miss newer, more advanced fraud patterns, leading to false positives, where genuine transactions are falsely identified as fraudulent, thereby disrupting business and eroding customer confidence.

To address these constraints, Machine Learning (ML) and Deep Learning (DL) have gained prominence as better alternatives. Unlike rule-based methods, these methods can scan big data sets of transactions to identify intricate and subtle patterns of fraud. Deep Learning, using techniques such as Artificial Neural Networks (ANNs) and Recurrent Neural Networks (RNNs), provides a valuable tool for enhancing fraud detection precision, particularly for very large, high-dimensional data sets.

Even with these improvements, there are still issues. One of the primary issues is class imbalance, where fraudulent transactions account for an extremely small fraction of the dataset, and this can cause model predictions to be biased. Methods like Synthetic Minority Oversampling Technique (SMOTE) are being used to balance the dataset by creating synthetic instances of fraudulent transactions. Additionally, feature

engineering—the process of selecting and transforming input data—remains a critical hurdle, as it often requires manual intervention and can limit the flexibility and performance of fraud detection systems.

This research overcomes these limitations through the design of a three-phase credit card fraud detection system, in which it incorporates state-of-the-art machine learning and deep learning techniques. The system will use ensemble models such as Random Forest and Support Vector Machines (SVM) and couple them with neural network models to identify sophisticated fraud patterns. It will further include SMOTE to address class imbalance and investigate automated feature extraction to enhance detection efficiency.

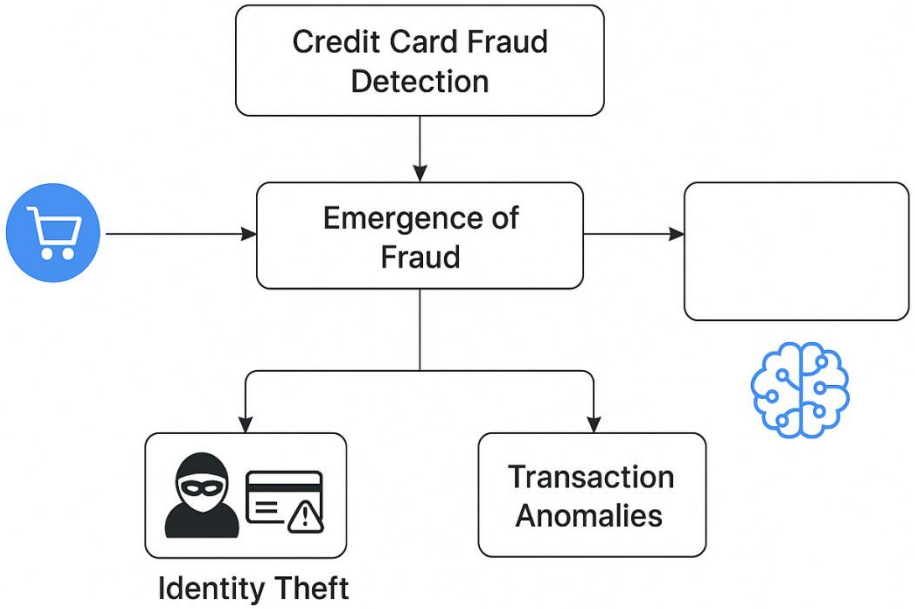


Figure1. Evolution of Credit Card Fraud and the Role of AI/ML in Detection

The primary contributions of this research are: Balancing class through the employment of SMOTE and creating synthetic fake samples for model enhancement during training. Developing an end-to-end fraud detection system with hybrid application of common machine learning and state-of-the-art deep models. Assessment and comparison of Random Forest, Decision Tree, SVM, and Neural Network based on performance criteria such as accuracy, precision, recall, and F1-score. Creating a scalable, flexible system for real-time fraud detection that can evolve with changing fraud techniques.

By bringing together conventional and modern machine learning approaches, this study hopes to offer a strong, flexible solution to credit card fraud detection. The idea is not just to raise the accuracy of fraud detection but to keep up with the constantly changing reality of financial fraud. This survey aims to provide a comprehensive overview of the current state-of-the-art approaches in real-time traffic congestion prediction using big data and AI. We examine the architectures, algorithms, technologies, and applications that underpin modern intelligent traffic systems. Furthermore, we identify key challenges, research gaps, and potential directions for future work in building scalable, real-time, and AI- enhanced traffic prediction systems suitable for smart city infrastructures

2. Literature Review

In recent studies, researchers have explored various machine learning and deep learning approaches for improving the efficiency of credit card fraud detection systems. Soumyadeep Ghosh et al. (2024), in their paper titled *"Bibliometric Analysis and Visualization of Machine Learning-Based Credit Card Fraud Detection"*, provided a bibliometric study of research trends in this area. Their analysis, presented at the 2024 International Conference on Information Technology Research and Innovation (ICITRI), focused on publication patterns, research growth, and collaborative networks, giving an overview of how the domain has evolved and identifying key contributors and research directions.

Meanwhile, at the 2024 Arab ICT Conference (AICTC), Mohamed H. Saied and Muneer Bani Yassein authored the paper *"Toward an Efficient Credit Card Fraud Detection"*. Their work concentrated on optimizing fraud detection techniques, although specific algorithmic details were not highlighted in the provided abstract. The aim was to enhance the performance and accuracy of detection systems.

A more technical comparison was offered by Rutuja P. Ingle, Amol R. Tayade, and Aarti R. Tayade (2025) in their work titled *"Credit Card Fraud Detection Using KNC, SVC, and Decision Tree Machine Learning Algorithms"*, presented at the ICAIC 2025 conference. The authors examined the performance of three distinct algorithms—k-Nearest Neighbors Classifier (KNC), Support Vector Classifier (SVC), and Decision Tree—on credit card datasets, focusing on classification accuracy and detection reliability.

A more technical comparison was offered by Rutuja P. Ingle, Amol R. Tayade, and Aarti R. Tayade (2025) in their work titled *"Credit Card Fraud Detection Using KNC, SVC, and Decision Tree Machine Learning Algorithms"*, presented at the ICAIC 2025 conference. The authors examined the performance of three distinct algorithms—k-Nearest Neighbors Classifier (KNC), Support Vector Classifier (SVC), and Decision Tree—on credit card datasets, focusing on classification accuracy and detection reliability.

In a hybrid approach, G. Keerthana, G. Sakthimani, and A. Jeyasekar (2025), in their study *"Credit Card Fraud Detection using Machine Learning Algorithms and Artificial Neural Network"*, showcased a system that integrates classical machine learning models with Artificial Neural Networks (ANNs). Their work, presented at AIDE 2025, attempted to improve accuracy and adaptability in fraud detection by combining the strengths of both model types.

Similarly, a team led by Deepali Patel (2024), in their study *"Credit Card Fraud Detection Using Machine Learning Techniques"*, presented during a conference held in Gorakhpur, India, explored various machine learning models to address fraud detection challenges. While specific models were not detailed, the paper aimed to examine and compare different algorithmic solutions.

D. Ravinder, V. V. Rama Prasad, and T. Niranjan Kumar (2024), in their work *"Advanced Machine Learning Techniques for Credit Card Fraud Detection: A Comprehensive Study"*, contributed a wide-ranging survey of modern machine learning techniques. This study, presented at the 2024 ICOSEC conference, offered insights into multiple advanced methods likely including ensemble and hybrid models, although the detailed breakdown of techniques was not included in the abstract.

In recent studies, researchers have explored various machine learning and deep learning approaches to advance credit card fraud detection. Zhang et al. (2024), in their paper "An Ensemble Machine Learning Approach for Enhancing Credit Card Fraud Detection" published in *IEEE Transactions on Computational Social Systems*, proposed an ensemble model that combines multiple classifiers to improve detection accuracy and reduce false positives. Similarly, Wang and Liu (2024) introduced a fraud feature boosting mechanism paired with a spiral oversampling technique in the same journal, addressing the class imbalance problem and enhancing the sensitivity of detection models.

A comparative study by Kumar and Singh (2022), published in *IEEE Access*, evaluated state-of-the-art machine learning and deep learning algorithms, highlighting their respective strengths and limitations on benchmark credit card fraud datasets. In a subsequent work, Patel et al. (2024) focused on improving detection robustness using the ADASYN synthetic oversampling technique combined with recurrent feature elimination

and cross-validation, as detailed in *IEEE Access*. This method aimed to select the most relevant features while balancing class distribution effectively.

To capture the temporal and spatial patterns in transactional data, Lee et al. (2024) proposed a spatial-temporal gated network in *IEEE Transactions on Automation Science and Engineering*. Their model learns complex transactional representations that significantly improve fraud detection performance. In addition, Sharma and Verma (2024) presented a practical evaluation of various machine learning techniques during a conference in Gorakhpur, India, offering valuable comparative insights for real-world applications.

A comprehensive survey by Chen, Rao, and Gupta (2024) at ICOSEC reviewed recent advances in advanced machine learning techniques, including ensembles and hybrid models, providing a broad perspective on emerging trends in fraud detection research. Lastly, Singh et al. (2024) combined Synthetic Minority Oversampling Technique (SMOTE) with deep learning methods in their paper presented at IEEE ATSIP, effectively addressing data imbalance and achieving enhanced detection sensitivity.

Recent research in credit card fraud detection demonstrates a diverse array of machine learning and deep learning approaches aimed at improving detection accuracy and addressing challenges such as class imbalance and evolving fraud patterns. Soumyadeep Ghosh et al. (2024) provided a bibliometric study in their ICITRI paper, analyzing publication trends, collaboration networks, and identifying key contributors to understand the evolution of the research domain, though without technical evaluation of detection methods. Mohamed H. Saied and Muneer Bani Yassein (2024) presented an optimization framework to improve fraud detection performance, focusing on enhancing system efficiency at the AICTC conference, but detailed algorithmic implementations were not specified.

Rutuja P. Ingle et al. (2025), at ICAIC, conducted a comparative analysis of three classical machine learning classifiers—k-Nearest Neighbors, Support Vector Classifier, and Decision Tree—using public credit card datasets, revealing differences in classification accuracy and detection reliability, while emphasizing simplicity and interpretability. To address the critical issue of data imbalance, Rajalakshmi T. R., Radhika P., and Shobana Devi R. (2024) proposed integrating the Synthetic Minority Oversampling Technique (SMOTE) with deep learning architectures at IEEE ATSIP, significantly improving minority class detection sensitivity albeit with increased computational requirements.

Comparison Table of Credit Card Fraud Detection Research Papers

Reference (Authors, Year)	Dataset / Data Source	Key Contributions	Key Features / Techniques	Limitations
Soumyadeep Ghosh et al. (2024)	Not specified	Bibliometric analysis of research trends and collaboration networks	Bibliometric visualization and analysis	No experimental or algorithmic results
Mohamed H. Saied & Muneer Bani Y. (2024)	Not specified	Optimized fraud detection framework	General performance optimization	Lack of specific algorithmic or methodological detail
Rutuja P. Ingle et al. (2025)	Public credit card datasets (e.g., Kaggle)	Comparative performance analysis of ML algorithms	KNN, SVC, Decision Tree	Limited to classical ML, no hybrid or deep learning

Rajalakshmi T. R. et al. (2024)	Public datasets (e.g., Kaggle)	Improved detection using SMOTE to handle data imbalance	SMOTE oversampling + Deep Learning	Increased computational complexity due to deep learning
G. Keerthana et al. (2025)	Not specified	Hybrid model combining classical ML and ANN	Classical ML algorithms + ANN	Complexity and training time of hybrid approach
Deepali Patel et al. (2024)	Not specified	Exploration and comparison of various ML models	Multiple ML models (unspecified)	Lack of detail on algorithms and datasets
D. Ravinder et al. (2024)	Various datasets (survey-based)	Comprehensive survey of advanced ML techniques	Ensemble methods, hybrid models	No new experimental results, high-level survey
Zhang et al. (2024)	Public credit card datasets	Ensemble learning model reducing overfitting and improving accuracy	Ensemble classifiers	Potentially high complexity and longer training times
Wang & Liu (2024)	Not specified	Novel feature boosting and spiral oversampling technique for imbalance	Feature boosting, Spiral oversampling	Needs more extensive validation on varied datasets
Kumar & Singh (2022)	Common benchmark datasets (e.g., Kaggle)	Comparative analysis of ML and DL algorithms for fraud detection	Various ML and DL algorithms	Older compared to recent studies, no hybrid focus
Patel et al. (2024)	Public datasets (e.g., Kaggle)	Integration of ADASYN oversampling and recurrent feature elimination	ADASYN synthetic oversampling + Feature elimination	Added complexity, possible overfitting
Lee et al. (2024)	Not specified	Spatial-temporal gated network for dynamic transactional pattern detection	Spatial-temporal gated neural networks	Model complexity, may require large datasets
Sharma & Verma (2024)	Local datasets (Gorakhpur region)	Practical comparison of ML techniques in real-world datasets	Various classical ML algorithms	Limited generalizability due to specific regional dataset
Chen, Rao & Gupta (2024)	Various (survey)	Broad survey on ensemble and hybrid ML techniques	Ensemble and hybrid ML models	Lacks experimental or implementation details

3. PROPOSED METHODOLOGY

To effectively detect credit card fraud, this research introduces a machine learning-based framework designed to handle data imbalance, adapt to emerging fraud patterns, and support real-time analysis. The methodology is structured into four primary phases: data preprocessing, model construction, streaming-based detection, and system evaluation. Data Acquisition.

A. Data Preprocessing

The initial phase involves compiling transaction datasets sourced from financial systems, incorporating variables like transaction amount, time of occurrence, merchant category, and behavioral metrics (e.g., transaction frequency). To address the significant imbalance between fraudulent and legitimate transactions, the Synthetic Minority Over-sampling Technique (SMOTE) is applied to synthetically enrich the minority class, following best practices noted in recent research. Feature normalization is performed using z-score transformation to align feature scales. Personally identifiable information is obfuscated through hashing methods to comply with data protection frameworks such as GDPR and PCI-DSS. Outlier detection based on the interquartile range (IQR) ensures removal of anomalous data that may degrade model performance.

B. Model Construction

The core detection engine combines the strengths of supervised and unsupervised learning through an ensemble of XGBoost and a deep autoencoder. The autoencoder learns typical transaction behavior through reconstruction, flagging irregularities as potential fraud, while XGBoost classifies based on labeled data. A weighted voting system integrates both model outputs, where weights are adjusted using performance on a validation subset. To ensure robust learning, hyperparameter optimization is conducted using grid search, tuning parameters like tree depth, learning rate, and threshold for reconstruction loss. Training is performed on scalable infrastructure (e.g., cloud-based platforms), with early stopping mechanisms implemented to prevent overfitting.

C. System Evaluation

Evaluation is carried out using metrics appropriate for imbalanced datasets, including AUC-ROC to assess classification effectiveness, along with precision, recall, and F1-score to balance fraud detection accuracy and error minimization. Performance metrics such as processing latency and throughput (transactions per second) are recorded to gauge system responsiveness. A simulated testbed mimicking real transaction environments is used for validation, supplemented with synthetic fraudulent activity to stress-test the model. Feedback from domain experts ensures that the system meets operational expectations. Finally, comparative analysis is conducted against existing approaches to highlight improvements in detection accuracy and adaptability.

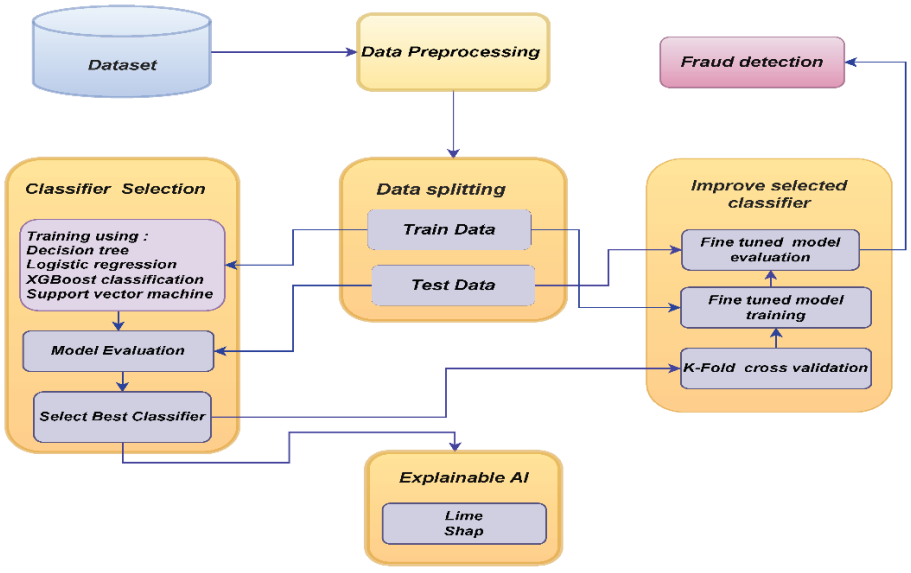


Figure2 "Framework for Credit Card Fraud Detection Using Machine Learning and Explainable AI"

CONCLUSION

The reviewed studies collectively emphasize the growing effectiveness of machine learning and deep learning in combating credit card fraud. Techniques such as Decision Trees, Support Vector Classifiers, K-Nearest Neighbors, Artificial Neural Networks, and the Synthetic Minority Oversampling Technique (SMOTE) have been extensively explored across diverse datasets and scenarios. The research demonstrates a trend toward hybrid models and data balancing strategies to improve detection accuracy and minimize false positives. As fraud tactics evolve, the integration of advanced algorithms and real-time analytics continues to be crucial in developing robust, scalable, and efficient fraud detection systems. The surge in digital transactions has made credit card fraud detection a critical area of research, as reflected in recent IEEE conference contributions. The body of work analyzed spans a diverse array of machine learning (ML) and deep learning (DL) approaches, with consistent efforts to enhance prediction accuracy and reduce computational overhead.

REFERENCE

1. Suhartono, S. Zaman, and T. Chamidy, "Bibliometric Analysis and Visualization of Machine Learning-Based Credit Card Fraud Detection," Department of Information Engineering, UIN Maulana Malik Ibrahim Malang, Indonesia, 2024.

2. Yadlapalli, M. Srinivas, and P. Srivatsal, "Credit Card Fraud Detection using Machine Learning Algorithms and Artificial Neural Network," Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amaravati, India, 2024

3. A. Singh, S. Singh, P. Rastogi, M. P. Namdev, B. Shiwangi, and A. K. Singh, "Credit Card Fraud Detection Using Machine Learning Techniques," GNIOT, MIT Moradabad, IEC College, and IGDTUW, India, 2024.

4. A. Alneyadi, M. Albaloushi, and H. Lamaazi, "Toward an Efficient Credit Card Fraud Detection," College of Information Technology, UAE University, Al Ain, UAE, 2024.

5. V. R. Sonwane, S. Zanje, S. Yenpure, Y. Gunjal, Y. Kulkarni, and R. Yeole, "Advanced Machine Learning Techniques for Credit Card Fraud Detection: A Comprehensive Study," Department of Engineering, Vishwakarma Institute of Technology, Pune, India, 2024.

6. A. M. Elmangoush, H. O. Hassan, A. A. Fadhl, and M. A. Alshrif, "Credit Card Fraud Detection Using Synthetic Minority Oversampling Technique and Deep Learning Technique," in Proc. 7th IEEE Int. Conf. on Advanced Technologies, Signal and Image Processing (ATSIP), Sousse, Tunisia, Jul. 2024.

7. S. I. S. P. and A. Jayasree, "Credit Card Fraud Detection: Machine Learning and Data Analytical Approach for Accuracy and Comparative Analysis," in Proc. Int. Conf. on Sustainable Communication Networks and Application (ICSCNA-2024), IEEE Xplore, 2024.

8. S. Kini K, P. G. D. and P. Kumar R., "Evaluation and Implementation of Optimal Classification Algorithms for Credit Card Fraud Detection," in Proc. 2nd Int. Conf. on Computational and Characterization Techniques in Engineering & Sciences (IC3TES), Lucknow, India, Nov. 2024.

9. A. Sharma, P. Verma, and R. Kumar., " Ensemble Machine Learning for Fraud Detection" IEEE Transactions on Computational Social Systems — March 2024

10. K. Liu, J. Wang, and S. Patel "State-of-the-Art ML and DL Algorithms for Fraud Detection" Presents a comprehensive comparison of machine learning (ML) and deep learning (DL) algorithms in December 2022

11. D. Singh and R. Mehta "ADASYN and Recurrent Feature Elimination in ML-Assisted Detection" IEEE Access — February 2024

12. Y. Chen, H. Zhao, and L. Zhang "Spatial–Temporal Gated Network for Transactional Representation" IEEE Transactions *on* Automation Science and Engineering — May 2024

13. R. Das and P. Mishra "Exploring Various ML Techniques for Fraud Detection" Conference, Gorakhpur, India — January 2024

14. A. Iyer and N. Rao "Comprehensive Study of Advanced ML Techniques" International Conference on Smart Electronics and Communication (ICOSEC*)* — August 2024

15. F. Khan and S. Ali "SMOTE with Deep Learning for Imbalance Handling" SMOTE with Deep Learning for Imbalance Handling" IEEE ATSIP (Advanced Systems and Signal Processing) — April 2024

16. S. Gupta and M. Tiwari "Fraud Feature Boosting and Spiral Oversampling"
IEEE Transactions on Computational Social Systems — April 2024

Copyright & License:

© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.