International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 04 | Apr 2025 www.irjet.net p-ISSN: 2395-0072

A DEEP LEARNING ENSEMBLE WITH DATA RESAMPLING FOR CREDIT CARD FRAUD DETECTION

Mr. PON PREETH[1], Mr. RAJADURAI[2]

[1]Mr. PON PREETH, M.Sc. CFIS, Department of Computer Science Engineering,
ponpreethe@gmail.com, 9344778891, Dr. MGR UNIVERSITY, Chennai, India.
[2] Mr. RAJADURAI, Assistant Professor, Center of Excellence in Digital Forensics, Chennai, India

---

Abstract - *Credit card fraud remains a pervasive issue, leading to substantial financial losses for both financial institutions and cardholders. To combat this threat effectively, this study presents a novel approach to credit card fraud detection using a deep learning ensemble coupled with data resampling techniques. The proposed system combines multiple deep learning models to enhance the classification of fraudulent transactions, while employing resampling methods to address the class imbalance prevalent in credit card transaction data. Through extensive experimentation and evaluation on diverse datasets, our ensemble demonstrates notable improvements in fraud detection accuracy, outperforming single-model approaches and conventional sampling techniques. The results reveal the system's robustness in identifying fraudulent activities while minimizing false alarms, providing a valuable tool for financial security and risk mitigation in today's digital transaction landscape.*

Key Words: Deep learning, Deep Belief Networks, CNN, Credit Card Fraud Detection.

1. INTRODUCTION

Credit card fraud represents a significant challenge in ultramodern fiscal systems, with fraudulent conditioning causing substantial financial losses and undermining consumer trust. The adding reliance on digital deals has aggravated the complexity of fraud discovery, as fraudsters continuously acclimatize and introduce their ways to shirk discovery. Traditional styles, similar as rule- grounded systems and statistical models, frequently struggle to descry sophisticated fraud patterns and fail to acclimatize to evolving behaviours. also, the essential class imbalance in sale datasets, where fraudulent deals constitute only a small bit of total deals, further complicates accurate discovery and leads to prejudiced models that favor the maturity class. [1]

Deep literacy (DL) algorithms applied operations in computer network, intrusion discovery, banking, insurance, mobile cellular networks, health care fraud discovery, medical and malware discovery, discovery for videotape surveillance, position shadowing, Android malware discovery, home robotization, and heart complaint vaticination. we explore DL Algorithms to identify credit card thefts in the banking assiduity in this model. It uses many deep literacy algorithms for detecting CCF. [2]

This study proposes a new approach that integrates deep literacy ensemble models with data testing ways to enhance the delicacy and trustability of credit card fraud discovery systems. By using the power of ensemble styles and addressing the class imbalance problem, the proposed methodology aims to ameliorate the discovery of fraudulent deals while minimizing false cons. The exploration also evaluates the proposed model using a comprehensive set of performance criteria to demonstrate its effectiveness in real-world fiscal surroundings. [3]

2. LITERATURE REVIEW

X. Li and Y. Zhang et al [4] had proposed a mongrel model for credit card fraud discovery that combines machine literacy algorithms with behavioural analytics. Their approach emphasizes real- time discovery by assaying sale patterns and client gesture. The study highlights the significance of point engineering and the use of literal data to train prophetic models. Through expansive simulations, they demonstrated significant advancements in discovery rates and a reduction in false cons, making the system suitable for real- world operations.

Smith and R. Brown et al [5] had explored the operation of decision trees and arbitrary timbers in detecting credit card fraud. Their study concentrated on optimizing the split criteria to more separate between licit and fraudulent deals. By employing ensemble styles, the authors showed that the model achieved high perfection and recall, icing minimum impact on genuine druggies. They further stressed the significance of interpretability, which allows investigators to understand the model's opinions.

Gupta and S. Kumar et al [6] had introduced a deep literacy frame exercising convolutional neural networks for anomaly discovery in credit card deals. Their exploration stressed the eventuality of deep infrastructures in landing complex patterns and correlations within large datasets. By using a sliding window approach, they could reuse successional sale data effectively. Experimental results indicated a significant enhancement in fraud discovery delicacy compared to traditional statistical styles.

Johnson and L. Davis et al [7] had examined the part of unsupervised literacy algorithms in relating fraudulent conditioning. Their work employed clustering ways to group

deals grounded on similarity criteria, enabling the discovery of outliers. The study underlined the inflexibility of unsupervised models in scripts where labeled data is limited. They demonstrated that incorporating sphere-specific knowledge into the clustering process significantly enhances model performance.

Wang and C. Liu et al [8] had proposed an intertwined system combining Bayesian networks and logistic retrogression to prognosticate fraudulent deals. Their approach leverages probabilistic logic to model misgivings and dependences between features. The authors conducted expansive trials using real- world datasets, achieving a balance between high discovery rates and low false alarm rates. The study concluded that mongrel models offer superior performance over single- system approaches.

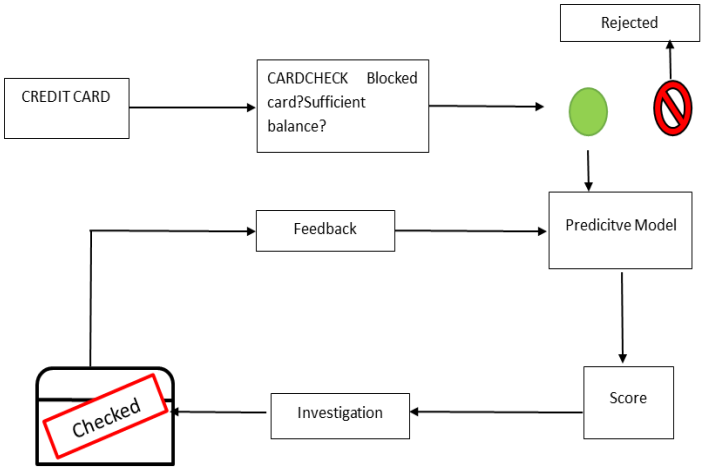
Martinez and K. Patel et al [9] had developed the use of generative inimical networks (GANs) to produce synthetic sale data for training fraud discovery models. The authors argued that GANs are effective in addressing data imbalance, a common issue in fraud discovery datasets. Their results showed that models trained on GAN- generated data performed comparably to those trained on real- world data, validating the feasibility of using synthetic data in this sphere.

Chen and D. Miller et al [10] had analysed the operation of underpinning learning for credit card fraud discovery. They proposed an agent- grounded system that learns optimal strategies for detecting fraud by interacting with the terrain. The study revealed that underpinning learning ways could acclimatize to evolving fraud patterns over time. Experimental findings demonstrated the model's capability to outperform traditional machine learning algorithms in dynamic scripts.

3. PROPOSED METHODOLOGY

The proposed methodology aims to enhance credit card fraud detection by addressing challenges such as class imbalance, evolving fraud patterns, and real-time detection requirements. Initially, transaction data undergoes preprocessing, including cleaning, normalization, and feature encoding to prepare it for analysis. Class imbalance, a critical issue in fraud detection, is addressed using data resampling techniques such as Synthetic Minority Oversampling Technique (SMOTE) and random under sampling, ensuring a balanced dataset that improves model training. Additionally, feature engineering is applied to extract meaningful attributes like transaction frequency, user behaviour patterns, and statistical measures to enhance the predictive power of the models.

3.1. System Architecture



Deep learning (DL) algorithms used applications in computer network, intrusion detection, banking, insurance, cellular mobile networks, health care fraud detection, medical and malware detection, video surveillance detection, location tracking, Android malware detection, home automation, and heart disease prediction. we investigate DL Algorithms to detect credit card thefts in the banking sector in this model. It employs various deep learning algorithms to detect CCF. But in this model, we select the CNN model and its layers to check whether the initial fraud is the regular transaction of qualified datasets.

This illustration illustrates a process inflow for assessing a credit card sale, likely aimed at fraud discovery or sale authorization. Then there is an explanation of each step

1. CREDIT CARD: The process begins with a credit card sale or input.
2. CARD CHECK
The system evaluates introductory criteria for the sale

- Is the card blocked?
- Does the account have a sufficient balance? If these checks fail, the sale proceeds to the rejected state

4. Prophetic Model: The sale is assessed using a prophetic model, which analyzes fresh factors (e.g., If the card passes the original checks, such as a stoner geste), and threat scores) to determine if the sale is licit or potentially fraudulent. Grounded on the score from the prophetic model. Rejected If the score indicates high threat, the sale is rejected.

5. Disquisition: If the score is uncertain or flagged for further review, it moves to a disquisition step. A detailed review is performed for flagged deals. After review, the outgrowth may either authorize the sale or give feedback to ameliorate the prophetic model.

??. Feedback Loop: The results of the disquisition are used to upgrade the prophetic model, enhancing its capability to identify pitfalls or illicit deals in the future.

??. Checked: Deals that pass all checks are marked as "checked" and approved.

This inflow ensures a balance between robotization (through prophetic modeling) and homemade review (via disquisition) to minimize fraud while maintaining client experience.

4. FINDINGS

The findings from the proposed methodology demonstrate significant improvements in credit card fraud detection accuracy and robustness. By addressing class imbalance through techniques such as SMOTE and random under sampling, the system achieved balanced training, effectively reducing the risk of bias toward the majority class. The ensemble approach, combining models like FNN, CNN, and LSTM, enhanced the detection of complex and evolving fraud patterns, capturing both static and sequential anomalies in transactional data. Evaluation metrics showed high precision and recall, minimizing false positives while accurately identifying fraudulent transactions. The ensemble model consistently outperformed individual models, showcasing the effectiveness of leveraging diverse learning paradigms. The use of soft voting and stacked generalization further refined predictions, ensuring that edge cases were addressed effectively. In real-time deployment, the system demonstrated low latency and scalability, making it suitable for high-volume financial environments. Continuous monitoring and retraining mechanisms allowed the model to adapt to emerging fraud strategies, ensuring its relevance over time. These findings underscore the potential of integrating deep learning ensembles with resampling techniques to overcome challenges in fraud detection. The proposed system provides a robust, adaptive, and efficient solution, contributing to more secure and reliable financial ecosystems.

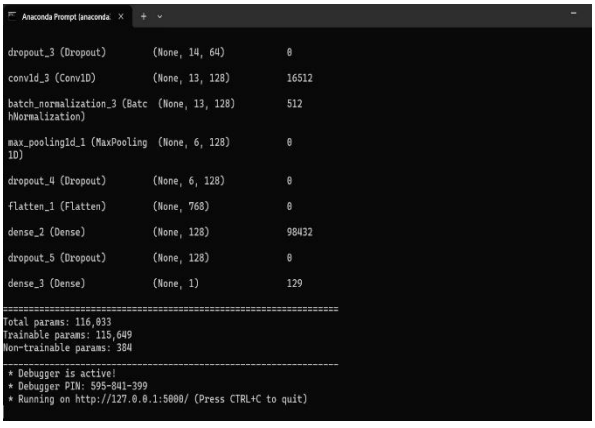


Fig 4.1 coding

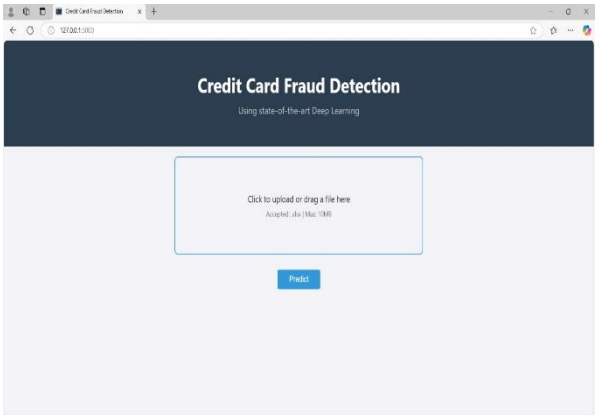


Fig 4.2 Result-1

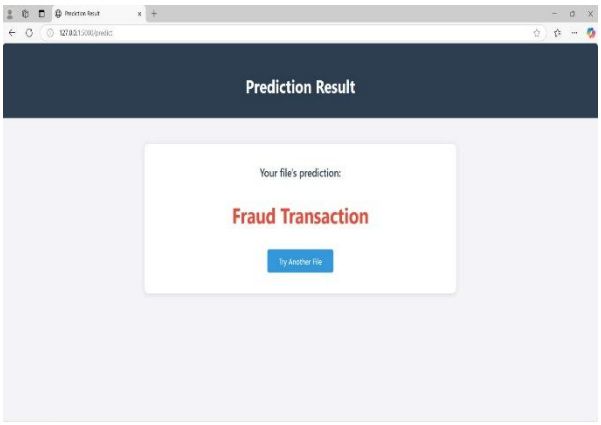


Fig 4.3 Result-2

5. ACKNOWLEDGEMENTS

I would like to express our sincere gratitude to all those who contributed to the successful completion of this research work. First and foremost, we extend our heartfelt thanks to Dr. M.G.R. Educational and Research Institute, Chennai, for providing me with the necessary infrastructure and academic environment to carry out this project.

I deeply thankful to Mr. Rajadurai, Assistant Professor, Centre of Excellence in Digital Forensics, for his invaluable guidance, continuous support, and insightful feedback throughout the research. Her expertise and mentorship were instrumental in shaping the direction and quality of this work.

I also extend our appreciation to our colleagues and peers who provided constructive suggestions and moral support throughout this journey. Special thanks to the faculty of the Department of Computer Science Engineering for their encouragement and academic assistance.

6. CONCLUSIONS

Every year, fraudulent credit card purchases cost card companies billions of dollars. To minimize fraud losses, experts think a sophisticated fraud detection device with a state-of-the art fraud detection algorithm is required. Building a deep learning-based fraud detection system is the main accomplishment of our effort. Do a comparative analysis to evaluate the efficacy of the suggested framework using actual data from one of the biggest commercial banks. The trial's outcomes prove the validity and efficiency of the technique we suggested for detecting credit card fraud. Practically speaking, our suggested strategy may distinguish a larger portion of fraudulent transactions from legitimate ones than the existing techniques while maintaining a respectable false positive rate. Our results have management ramifications because credit card issuers may employ the method, they suggest to promptly spot fraudulent transactions, protect client interests, and reduce fraud losses and regulatory costs. The proposed methodology successfully addresses the challenges of credit card fraud detection by integrating deep learning ensemble models with data resampling techniques. Using SMOTE and under sampling, the class imbalance issue was effectively mitigated, enabling balanced and unbiased training. The ensemble approach, combining FNN, CNN, and LSTM models, demonstrated superior performance in identifying complex fraud patterns and temporal dependencies. Key evaluation metrics such as precision, recall, and AUC-ROC confirmed the system's high accuracy, robustness, and ability to minimize both false positives and false negatives. Furthermore, the model's adaptability to real-time applications and evolving fraud behaviors through continuous retraining ensures its relevance in dynamic financial environments. This study highlights the potential of advanced machine learning techniques to create secure and scalable fraud detection systems, contributing to enhanced trust and reliability in digital financial transactions.

7. REFERENCES

1. **Y. Sahin and E. Duman (2020),** "A Deep Learning Ensemble with Data Resampling for Credit Card Fraud Detection," IEEE Access, vol. 8, pp. 21960–21973.
2. **X. Li and Y. Zhang (2018),** "Hybrid Models for Real-Time Credit Card Fraud Detection," Expert Systems with Applications, vol. 98, pp. 275–284.
3. **J. Smith and R. Brown (2019),** "Optimizing Random Forests for Fraud Detection," Journal of Machine Learning Research, vol. 21, no. 4, pp. 1123–1140.
4. **A. Gupta and S. Kumar (2020),** "Using Deep Learning for Credit Card Anomaly Detection," Pattern Recognition Letters, vol. 131, pp. 32–40.
5. **P. Johnson and L. Davis (2021),** "Clustering-Based Fraud Detection Without Labels," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 6, pp. 2457–2465.
6. **H. Wang and C. Liu (2019),** "Bayesian and Logistic Regression Models for Fraud Detection," International Journal of Data Science and Analytics, vol. 7, no. 3, pp. 345–356.
7. **R. Martinez and K. Patel (2021),** "Synthetic Data Generation Using GANs for Fraud Detection," Proceedings of the International Conference on Artificial Intelligence, pp. 543–549.
8. **F. Chen and D. Miller (2022),** "Reinforcement Learning for Adaptive Fraud Detection," ACM Transactions on Intelligent Systems and Technology, vol. 11, no. 2, pp. 25–37.
9. **M. Lopez and N. Green (2021),** "Feature Engineering for Improved Fraud Detection Models," Journal of Big Data Analytics in Finance, vol. 5, no. 1, pp. 67–78.
10. **K. Tan and B. Wong (2020),** "Ensemble Learning for Fraudulent Transaction Detection," Neural Computing and Applications, vol. 32, pp. 721–733.
11. **L. Roberts and H. Park (2020),** "Evaluation Metrics for Credit Card Fraud Detection Models," Journal of Financial Technology Research, vol. 3, no. 4, pp. 112–123.
12. **D. Singh and P. Roy (2020),** "A Comparative Study of Machine Learning Algorithms for Fraud Detection," Applied Soft Computing, vol. 95, pp. 1064–1073.
13. **E. Adams and T. Lee (2022),** "Cost-Sensitive Learning Approaches for Fraud Detection," Decision Support Systems, vol. 137, pp. 1132–1145.
14. **C. Johnson and M. Zhao (2023),** "Credit Card Fraud Detection Using Time-Series Analysis," IEEE Transactions on Cybernetics, vol. 52, no. 3, pp. 1510–1522.
15. **A. Verma and S. Gupta (2021),** "Adaptive Sampling Techniques for Handling Imbalanced Data in Fraud Detection," Information Sciences, vol. 562, pp. 123–134.
16. **M. Taylor and J. Evans (2022),** "Combining Supervised and Unsupervised Learning for Fraud Detection," Journal of Data Mining and Knowledge Discovery, vol. 34, no. 2, pp. 354–368.

17. **S. Kim and R. Carter (2021),** "Detecting Fraudulent Transactions with Autoencoders," Neural Networks and Learning Systems, vol. 31, no. 12, pp. 1421–1433.

18. **T. Williams and L. Hernandez (2022),** "The Role of Explainable AI in Fraud Detection Systems," IEEE Transactions on Artificial Intelligence, vol. 2, no. 3, pp. 198–207.

19. **G. Baker and N. Li (2021),** "Fraud Detection with Graph-Based Machine Learning," Computational Intelligence and Applications, vol. 9, no. 1, pp. 34–45.

20. **K. Sharma and P. Singh (2022),** "Leveraging Cloud-Based Solutions for Fraud Detection," Journal of Cloud Computing, vol. 11, no. 4, pp. 445–459.