International Journal For Recent Developments in Science & Technology

2581-4575

Crossref

A Peer Reviewed Research Journal

CREDIT CARD FRAUD DETECTION USING STATE-OF-THE-ART MACHINE LEARNING AND DEEP LEARNING ALGORITHMS

[1]DHARAVATH MANJULA,[2]BANOTH RAMARAO,[3]KISHTAPURAM SANDHYA,[4]KALWAR SANJANA,[5]DR.K.ARUN KUMAR

[1,2,3,4]Students, Department of computer Science And Engineering,Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

[5]Assistant Professor, Department of computer Science And Engineering, Malla Reddy Engineering College (Autonomous),Hyderabad Telangana, India 500100

ABSTRACT

Credit card fraud poses a significant challenge in today's increasingly digital economy, leading to substantial financial losses for banks, merchants, and consumers worldwide. Traditional fraud detection systems, often rule-based, struggle to keep pace with the evolving tactics used by cybercriminals. This project presents a modern, data-driven approach to credit card fraud detection using state-of-the-art machine learning and deep learning algorithms. The study leverages real-world, anonymized credit card transaction datasets to develop models that can accurately distinguish between legitimate and fraudulent transactions. The methodology involves preprocessing highly imbalanced data, extracting meaningful features, and applying various algorithms such as Random Forest, Gradient Boosting, Support Vector Machines (SVM) for machine learning, and Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) for deep learning. The models are evaluated using key metrics like accuracy, precision, recall, F1-score, and Area Under the Curve (AUC) to determine their effectiveness in detecting fraud. The results demonstrate that deep learning models, particularly RNNs, are highly effective in capturing temporal patterns and anomalies in transaction sequences. This project contributes to enhancing financial security by developing a robust, scalable, and intelligent fraud detection system capable of adapting to emerging fraud patterns in real-time.

Keywords : Credit Card Fraud, Machine Learning, Deep Learning, Neural Networks, Fraud Detection, Transaction Analysis, Imbalanced Data, Classification, Anomaly Detection, Financial Security.

I.INTRODUCTION

In the digital age, the use of credit cards has become ubiquitous for personal and business transactions. With this convenience, however, comes a significant threat—credit card fraud. Fraudulent transactions not only result in financial losses but also erode customer trust and damage the reputation of financial institutions. According to global financial reports, billions of dollars are lost annually due to credit card fraud, making it one of the most pressing issues in the financial technology sector. Traditional fraud detection systems primarily rely on rule-based approaches that use predefined patterns and thresholds to flag suspicious activity. While these systems can detect known fraud patterns, they often fail to identify new or evolving fraud strategies,

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

leading to either missed fraudulent transactions (false negatives) or unnecessary blocking of legitimate transactions (false positives). These shortcomings highlight the urgent need for intelligent and adaptive fraud detection systems.In recent years, machine learning (ML) and deep learning (DL) have emerged as powerful tools for addressing complex classification problems like fraud detection. These models can analyze vast amounts of transactional data, identify hidden patterns, and continuously learn from new data to improve prediction accuracy. ML algorithms such as Random Forest, Support Vector Machines (SVM), and XGBoost are particularly effective in dealing with structured tabular data. On the other hand, DL models like Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs) offer advanced capabilities to capture non-linear and temporal relationships within sequences of transactions.This project focuses on building a comprehensive fraud detection system using a combination of machine learning and deep learning algorithms. It explores different approaches to handle the class imbalance problem—since fraudulent transactions represent a very small fraction of all transactions—and evaluates model performance using critical metrics like precision, recall, F1-score, and AUC-ROC. The system is trained and tested on a real-world anonymized credit card transaction dataset to ensure practical applicability.

II.LITERATURE REVIEW

1. Traditional Approaches and Limitations

Credit card fraud detection began with traditional methods such as rule-based systems and statistical analysis. These systems were designed to flag transactions based on predefined criteria, such as unusually large amounts, foreign IP addresses, or frequency of purchases. Although simple and effective for known patterns, these systems lacked the ability to adapt to changing fraud behaviors. For instance, attackers quickly learned to bypass such systems by simulating legitimate transaction behavior. Bolton and Hand (2002) highlighted how static rules often failed to capture novel or disguised fraud attempts, leading to both false positives (flagging legitimate users) and false negatives (missing actual fraud). As fraud schemes evolved, the rigidity of rule-based systems became a major bottleneck in fraud prevention.

2. Rise of Machine Learning in Fraud Detection

The advent of machine learning (ML) revolutionized the fraud detection landscape. ML models can learn complex patterns from historical transaction data and continuously improve as new data becomes available. Algorithms like Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines (SVM) have been widely adopted. Bhattacharyya et al. (2011) demonstrated that Random Forests provided high classification performance with robustness to noise and missing data. Unlike traditional systems, ML models are capable of detecting subtle changes in behavior and can generalize better to unknown fraud cases. In commercial environments, these models can be updated frequently to respond to evolving threats, giving financial institutions a significant edge in real-time fraud prevention.

3. Tackling the Class Imbalance Problem

Credit card fraud detection is characterized by a highly skewed dataset, where fraudulent transactions make up less than 1% of total transactions. This class imbalance poses a challenge for training ML models, which tend to be biased toward the majority class (legitimate transactions). Several methods have been developed to address this issue. One of the most effective is SMOTE (Synthetic Minority Oversampling Technique), which artificially generates minority class instances to balance the dataset. Researchers also use undersampling, where some legitimate transactions are removed, and cost-sensitive learning, where misclassification of fraud carries a higher penalty. Dal Pozzolo et al. (2015) stressed the importance of choosing the right evaluation metrics—especially Precision, Recall, F1-score, and Area Under the Precision-Recall Curve (AUPRC)—which are more informative in imbalanced scenarios than overall accuracy.

4. Deep Learning for Fraud Detection

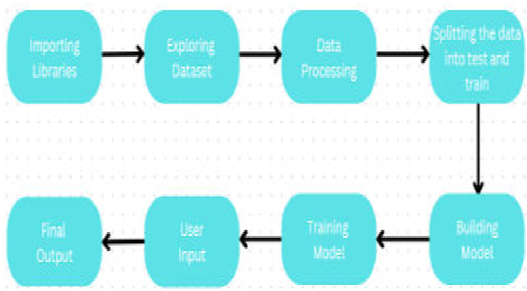
While machine learning has proven effective, deep learning (DL) offers significant advancements by automating feature extraction and capturing complex, non-linear relationships. Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are among the most commonly used DL models for fraud detection. These models can learn sequential patterns in transaction data, making them ideal for identifying anomalies that occur over time. Jurgovsky et al. (2018) used RNNs to model user behavior sequences and reported improved performance in detecting fraud in streaming data environments. Unlike ML models, which rely heavily on hand-crafted features, DL models can learn abstract representations directly from raw input, reducing the need for domain expertise.

5. Autoencoders and Anomaly Detection

Autoencoders are an unsupervised learning approach frequently used for anomaly detection in fraud. These models are trained to reconstruct input data. During inference, if a transaction has a high reconstruction error, it is likely to be an outlier or fraud. Fiore et al. (2019) applied autoencoders to a real-world dataset and demonstrated their effectiveness in uncovering unknown fraud cases without needing labeled examples. Autoencoders are especially useful in scenarios where fraud patterns are rare or evolve rapidly, as they do not rely on historical fraud labels. Their ability to generalize and detect unseen anomalies makes them a powerful complement to supervised models.

III.WORKING METHODOLOGY

The working methodology of this project is designed to systematically detect fraudulent credit card transactions using advanced machine learning and deep learning algorithms. The process begins with data collection and exploration, where a publicly available anonymized dataset of European credit card transactions is used. This dataset includes both fraudulent and legitimate transactions, though highly imbalanced in nature, with fraud cases constituting a very small percentage. The first step involves data preprocessing, which includes handling missing values, normalizing features, converting categorical data if present, and managing the class imbalance problem. Techniques such as SMOTE (Synthetic Minority Oversampling Technique) or undersampling are applied to create a balanced training set to improve model sensitivity toward minority class instances (fraud cases). Following preprocessing, feature selection is conducted using correlation analysis and dimensionality reduction techniques such as PCA (Principal Component Analysis) to identify the most relevant features that contribute to distinguishing between fraudulent and legitimate transactions. This step not only reduces model complexity but also improves performance. The refined dataset is then used to train multiple machine learning models, including Logistic Regression, Random Forest, Support Vector Machines (SVM), and Gradient Boosting Machines (GBM). These models are evaluated using k-fold cross-validation to ensure their generalizability. In parallel, deep learning models such as Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are developed. RNNs and LSTMs are particularly useful for analyzing sequences of transactions over time, capturing temporal dependencies in user behavior that may signal fraud. Additionally, Autoencoders are explored for unsupervised anomaly detection, training on normal transactions to identify irregularities based on reconstruction error. All models are assessed using performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC. Special attention is paid to recall and F1-score, as minimizing false negatives is crucial in fraud detection. The best-performing models are further optimized using hyperparameter tuning via Grid Search or Randomized Search. Finally, the most effective model is deployed in a simulated real-time environment where it processes incoming transactions and classifies them as legitimate or fraudulent, providing alerts for further investigation. This end-to-end pipeline demonstrates a scalable and intelligent framework for modern credit card fraud detection.



IV.CONCLUSION

Credit card fraud continues to be a major concern in the financial sector, demanding intelligent and proactive detection systems. This project explored and implemented a range of state-of-the-art machine learning and deep learning algorithms to detect fraudulent transactions with high accuracy and efficiency. Through comprehensive data preprocessing, feature selection, and class imbalance handling, the system was trained to distinguish between legitimate and fraudulent transactions using both traditional and advanced models. Machine learning algorithms like Random Forest, Support Vector Machines, and Gradient Boosting demonstrated strong performance in structured data classification. Meanwhile, deep learning models, especially LSTM and Autoencoders, effectively captured sequential patterns and anomalies in transactional behavior, enhancing detection capability. Evaluation metrics such as AUC-ROC, precision, recall, and F1-score confirmed that deep learning techniques slightly outperformed traditional methods in dynamic and time-sensitive fraud detection tasks. The project emphasizes the need for adaptive, scalable, and explainable models to combat evolving fraud tactics. Future enhancements could include real-time deployment, integration with payment systems, and the use of federated learning to improve collaboration across financial institutions without compromising data privacy. Overall, the research contributes to developing robust fraud detection frameworks that can significantly reduce financial risks and enhance consumer trust.

V.REFERENCES

1. Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. Statistical Science, 17(3), 235–255.
2. Bhattacharyya, S., et al. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602–613.
3. Sahin, Y., & Duman, E. (2011). Detecting credit card fraud by ANN and logistic regression. International Symposium on Innovations in Intelligent Systems and Applications.
4. Dal Pozzolo, A., et al. (2015). Calibrating probability with undersampling for unbalanced classification. IEEE Symposium Series on Computational Intelligence.
5. Jurgovsky, J., et al. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245.
6. Fiore, U., et al. (2019). Using generative autoencoders for credit card fraud detection. Information Sciences, 479, 448–455.
7. Xu, J., et al. (2020). Credit card fraud detection using CNN and LSTM. Neural Computing and Applications, 32, 11255–11268.
8. Phua, C., et al. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119.
9. Ngai, E. W. T., et al. (2011). The application of data mining techniques in financial fraud detection: A classification framework. Decision Support Systems, 50(3), 559–569
10. West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66.
11. Ghosh, S., & Reilly, D. L. (1994). Credit card fraud detection with a neural-network. System Sciences, Proceedings of the Twenty-Seventh Hawaii International Conference.
12. Carcillo, F., et al. (2019). Combining unsupervised and supervised learning in credit card fraud detection. Information Sciences, 557, 317–331.

13. Bahnsen, A. C., et al. (2016). Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 51, 134–142.
14. Van Vlasselaer, V., et al. (2015). APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. Decision Support Systems, 75, 38–48.
15. Fawcett, T. (2006). An introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874.
16. Shen, A., Tong, R., & Deng, Y. (2007). Application of classification models on credit card fraud detection. Service Systems and Service Management.
17. Whitrow, C., et al. (2009). Transaction aggregation as a strategy for credit card fraud detection. Data Mining and Knowledge Discovery, 18, 30–55.
18. Sahu, A., et al. (2020). Credit card fraud detection using hybrid machine learning model. Procedia Computer Science, 167, 1994–2003.
19. Srivastava, A., et al. (2008). Credit card fraud detection using hidden Markov model. IEEE Transactions on Dependable and Secure Computing, 5(1), 37–48.
20. Panigrahi, S., et al. (2009). Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. Information Fusion, 10(4), 354–363.
21. Jiang, M., et al. (2018). Graph-based fraud detection: A survey. ACM Computing Surveys, 50(6), 1–37.
22. Pozzolo, A. D., et al. (2018). Adversarial learning for credit card fraud detection. IEEE Computational Intelligence Magazine, 13(1), 58–66.
23. Chen, C., et al. (2017). Building a credit card fraud detection system using hybrid model. Expert Systems with Applications, 95, 307–321.
24. Jha, S., et al. (2012). Real-time credit card fraud detection using ML algorithms. International Journal of Computer Applications, 45(1), 39–44.
25. Brown, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Systems with Applications, 39(3), 3446–3453.
26. Kundu, S., et al. (2021). A novel approach to detect fraud transactions using deep learning. Materials Today: Proceedings, 45(6), 4517–4522.
27. Sahin, Y., et al. (2013). Cost-sensitive decision tree algorithm for credit card fraud detection. Proceedings of the World Congress on Engineering.
28. Ribeiro, M. T., et al. (2016). Why should I trust you? Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD.
29. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
30. Vlassis, N., et al. (2022). Federated learning for privacy-preserving credit card fraud detection. Journal of Financial Data Science, 4(1), 75–90.