JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Credit Card Fraud Detection Using Machine Learning Approaches

Shruti Pacharne , Rupina Nadar, Payal Ugale, H.R.Kulkarni , S. H. Karande, Amit More*
G H raisoni College of Arts,Commerce And Science,Wagholi,Maharashta,India
*-Author For Correspondence. Email: amit2000bm@gmail.com

Abstract

The rise of digital transactions has increased both convenience and financial risks, particularly through credit card fraud. As fraudsters adopt advanced techniques, traditional rule-based systems are no longer sufficient for effective detection. This research presents a machine learning driven framework designed to identify fraudulent transactions with high accuracy and minimal false alarms. Several supervised, unsupervised, and hybrid algorithms are analyzed to evaluate their performance on imbalanced financial datasets. The results highlight that blended learning approaches offer superior detection capabilities while remaining scalable for real-time environments. The study aims to support financial institutions and government bodies in strengthening digital payment security.

Keywords

Credit Card Fraud, Data Mining, Machine Learning, Anomaly Detection, Financial Security, Predictive Modelling

1. Introduction

Digital payments and credit card usage have grown rapidly due to online shopping, electronic banking, and mobile-based transactions. Alongside this growth, credit card fraud has become a major challenge for customers, merchants, and financial institutions. Fraudulent activities include unauthorized transactions, identity theft, card cloning, and false chargebacks.

Conventional fraud detection rules such as transaction limits, geographic restrictions, and merchant validation are static and ineffective against new fraud strategies. Machine learning, however, enables systems to learn from real data and adapt to evolving patterns. This paper explores how machine learning can enhance fraud detection and proposes a reliable detection framework suitable for government and banking environments.

2. Literature Review (Related Work)

2.1 Early Fraud Detection Approaches

Early systems relied heavily on manual or rule-based mechanisms. Although simple, such systems struggle to detect complex and previously unseen fraud patterns.

2.2 Supervised Learning Techniques

Methods such as Logistic Regression, Decision Trees, Random Forest, SVM, and Gradient Boosting have shown strong results on labeled transaction data. These models identify fraud based on past examples but depend on high-quality labeled datasets.

2.3 Unsupervised Learning Methods

Anomaly-based algorithms like Isolation Forest, K-Means, and Autoencoders detect deviations in customer behavior. These models are effective for discovering new fraud trends but may generate more false positives.

2.4 Hybrid Approaches

Researchers have combined supervised and unsupervised models to improve accuracy. For instance, anomaly detection is used for initial filtering, and a classifier is used for final verification. This dual-stage approach reduces noise and strengthens prediction.

2.5 Deep Learning Models

Advanced neural architectures such as LSTMs for sequential transactions and deep autoencoders have demonstrated high detection capabilities by learning complex transaction patterns.

Overall, literature indicates that hybrid and deep learning models provide the most robust performance in modern fraud detection.

3. Problem Definition

Credit card fraud detection involves several key challenges:

* Severely Imbalanced Data – Fraud transactions are extremely rare compared to normal ones.

* Dynamic Fraud Behavior – Fraudsters frequently adopt new techniques to bypass security systems.

* Real-Time Decision Requirements – Financial institutions must detect fraud within milliseconds.

* High Volume of Data – Banking systems process millions of transactions daily.

* Customer Impact – False alerts create inconvenience and damage institutional trust.

The research aims to build a system capable of accurate, fast, and adaptable fraud detection.

4. Methodology

4.1 Data Collection

An anonymized dataset containing transaction amount, merchant details, time stamps, customer patterns, and fraud labels is used for experimentation.

4.2 Data Preprocessing

* Removal of missing and inconsistent entries

* Scaling numerical attributes using standardization

* Encoding categorical values

* Balancing the dataset using SMOTE to reduce class imbalance

4.3 Feature Engineering

Key features derived include:

* Transaction frequency per customer

* Sudden deviation in spending pattern

* Location inconsistencies

* Merchant risk categorization

* Time-based behavior analysis

4.4 Model Selection

Three groups of algorithms were tested: Supervised:
Random Forest, Logistic Regression, XGBoost, SVM Unsupervised:
Isolation Forest, Autoencoders Hybrid:
Autoencoder followed by Random Forest classifier

4.5 Evaluation Metrics

Models were assessed using accuracy, precision, recall, F1-score, and ROC-AUC to ensure balanced performance.

5. Implementation

5.1 System Architecture

The proposed architecture includes:

* Data Layer – Handles storage and preprocessing

* Model Layer – Manages training and prediction

* Fraud Detection Layer – Evaluates transaction risk in real time

* Alert Layer – Generates notifications for high-risk cases

5.2 Technology Stack

* Python

* Scikit-Learn, TensorFlow/Keras

* Pandas, NumPy

* PostgreSQL database

* REST APIs for integration with financial services

5.3 Operational Flow

* Transaction enters the detection pipeline.

* Model calculates fraud probability.

* If probability exceeds threshold → flag as suspicious.

* Alert sent to fraud investigation team.

* Case stored for model retraining and improvement.

6. Results and Discussion

6.1 Insights

* Gradient boosting and tree-based models perform strongly on structured datasets.

* Hybrid models excel because they detect both known and unknown fraud cases.

* High precision reduces false alarms important for customer satisfaction.

* Models are capable of real-time processing, suitable for financial environments.

6.2 Applicability

This framework can be implemented by:

* Government financial oversight systems

* Banks and credit card networks

* FinTech and digital wallet organizations

* Cybersecurity monitoring platforms

7. Conclusion

Credit card fraud continues to threaten financial security, especially in rapidly expanding digital economies. Machine learning introduces advanced capabilities to detect unusual patterns and predict fraudulent behavior more efficiently than traditional methods. This research demonstrates that hybrid machine learning models achieve the highest performance, offering strong accuracy and adaptability. The proposed framework can be integrated into large-scale banking and government systems to enhance the safety of digital transactions and reduce economic losses.

8. References

[1] Jiang, Changjun et al. "Credit Card Fraud Detection: A Novel Approach Using Aggregation Strategy and Feedback Mechanism." IEEE Internet of Things Journal 5 (2018): 3637-3647.

[2]Pumsirirat, A. and Yan, L. (2018). Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. International Journal of Advanced Computer Science and Applications, 9(1).

[3]Mohammed, Emad, and Behrouz Far. "Supervised Machine Learning Algorithms for Credit Card Fraudulent Transaction Detection: A Comparative Study." IEEE Annals of the History of Computing, IEEE, 1 July 2018, doi.ieeecomputersociety.org/10.1109/IRI.2018.00025.

[4]Randhawa, Kuldeep, et al. "Credit Card Fraud Detection Using AdaBoost and Majority Voting." IEEE Access, vol. 6, 2018, pp. 14277–14284., doi:10.1109/access.2018.2806420.

[5]Roy, Abhimanyu, et al. "Deep Learning Detecting Fraud in Credit Card Transactions." 2018 Systems and Information Engineering Design Symposium (SIEDS), 2018, doi:10.1109/sieds.2018.8374722.

[6]Xuan, Shiyang, et al. "Random Forest for Credit Card Fraud Detection." 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018, doi:10.1109/icnsc.2018.8361343.

[7]Awoyemi, John O., et al. "Credit Card Fraud Detection Using Machine Learning Techniques: A Comparative Analysis." 2017 International Conference on Computing Networking and Informatics (ICCNI), 2017, doi:10.1109/iccni.2017.8123782.

[8]Melo-Acosta, German E., et al. "Fraud Detection in Big Data Using Supervised and Semi-Supervised Learning Techniques." 2017 IEEE Colombian Conference on Communications and Computing (COLCOM), 2017, doi:10.1109/colcomcon.2017.8088206.

[9] http://www.rbi.org.in/Circular/CreditCard

[10] https://www.ftc.gov/news-events/press-releases/2019/02/imposter-scams-top-complaints-made-ftc-2018

[11] https://www.kaggle.com/mlg-ulb/creditcardfraud

[12] https://www.kaggle.com/uciml/default-of-credit-card-clients-dataset