An Analytical Study on AI and Machine Learning Techniques for Financial Fraud Detection

Names: Veda Jadhav, Disha Jagtap
Course: M.Sc. (Data Science) Sem-III, Savitribai Phule Pune University

Abstract

This research analyzes the application of Artificial Intelligence (AI) and Machine Learning (ML) models in detecting financial fraud in digital transactions. Traditional rule-based systems often fail to adapt to evolving fraud patterns. The study focuses on comparing supervised, unsupervised, and hybrid models to identify those that provide higher detection accuracy and adaptability. Algorithms such as Logistic Regression, Random Forest, XGBoost, Autoencoders, and LSTM are evaluated using performance metrics like accuracy, precision, recall, and F1-score. Findings indicate that ensemble and deep learning approaches significantly enhance detection performance while maintaining real-time scalability.

Introduction and Background

With the rapid growth of digital banking, credit cards, and e-commerce, financial fraud has become more sophisticated and difficult to detect. Conventional rule-based systems rely on static thresholds and manual reviews, making them slow and inefficient. AI and ML algorithms, however, learn from historical data to identify unusual behavioral or transactional patterns. The research aims to evaluate multiple ML algorithms that can automatically detect anomalies and reduce false positives in large-scale financial datasets.

Research Problem and Objectives

Problem Statement: Existing fraud detection systems suffer from high false-positive rates and poor adaptability to new fraud techniques. They lack interpretability and fail to process large volumes of financial data in real time.
Objectives:

- To compare supervised, unsupervised, and hybrid AI/ML models for fraud detection.
- To evaluate models based on accuracy, precision, recall, and F1-score.
- To design an adaptive framework integrating explainable AI (XAI) for transparency.
- To identify implementation challenges and propose solutions for scalable fraud detection.

Literature Review

Recent research demonstrates the growing efficiency of ensemble and deep learning models in financial fraud detection. Bhattacharya et al. (2021) highlighted that ensemble models like Random Forest outperform single classifiers. Carcillo et al. (2019) showed that combining supervised and unsupervised techniques enhances robustness. Ghosh et al. (2022) addressed the need for explainable AI to overcome transparency issues in deep neural networks. Hybrid frameworks that use clustering followed by classification (Zhang et al., 2021) have proven effective in minimizing noise and improving accuracy.

Research Methodology

The study used financial transaction data, both real and simulated, containing attributes like transaction amount, merchant type, and class labels (fraud/non-fraud). Data preprocessing involved normalization, one-hot encoding, and handling imbalance using SMOTE. Algorithms such as Logistic Regression, Random Forest, XGBoost, Autoencoder, and LSTM were implemented in Python using Scikit-learn and TensorFlow. Models were evaluated using 10-fold cross-validation, and metrics like accuracy, precision, recall, and F1-score were computed.

Results and Discussion

Experimental results show that XGBoost achieved the highest accuracy (95.6%) and F1-score (0.93), followed by Random Forest (94.2%). Deep learning models like LSTM demonstrated strong adaptability but required higher computational resources. Autoencoders effectively detected anomalies but lacked interpretability. The comparative results confirm that ensemble methods balance accuracy, scalability, and explainability more effectively than other models.

Expected Outcomes and Future Scope

The research is expected to aid financial institutions in developing intelligent fraud detection frameworks capable of real-time learning and adaptation. Future work can explore Explainable AI (XAI) for regulatory transparency, Federated Learning for privacy-preserving training, and Blockchain integration for transaction integrity. Additionally, Graph Neural Networks (GNNs) can help uncover hidden fraud networks by analyzing relational transaction data.

References

Bhattacharya, S., et al. (2021). Machine Learning Approaches for Financial Fraud Detection. Journal of Financial Analytics.
Carcillo, F., et al. (2019). Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection. IEEE Transactions on Neural Networks.
Ghosh, R., et al. (2022). Deep Learning in Fraud Detection: Trends and Challenges. Int. Journal of AI Applications.
Nguyen, D., & Tran, M. (2020). Ensemble Learning Models for Financial Fraud Detection. Springer Open Access.
Zhang, W., et al. (2021). Hybrid Machine Learning Models for Credit Card Fraud Detection. Expert Systems with Applications.