IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 11, November 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.524

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024.1311090|

AI-Driven Techniques for Fraud Detection in Credit Card Transactions

Vishnu C S, Zaiba Elmi, Navanidhi S Shetty, Suhas K C

U.G. Student, Department of Computer Engineering,Channabasaveshwara Institute of Technology, Tumakuru, Karnataka, India

U.G. Student, Department of Computer Engineering, Channabasaveshwara Institute of Technology, Tumakuru, Karnataka, India

U.G. Student, Department of Computer Engineering, Channabasaveshwara Institute of Technology, Tumakuru, Karnataka, India

Associate Professor, Department of Computer Engineering, Channabasaveshwara Institute of Technology ,Tumakuru, Karnataka, India

ABSTRACT:Credit card fraud remains an important issue facing financial institutions, leading to extensive financial losses and a decrease in public confidence in payment systems. Traditional fraud detection methods, often based on fixed rule sets, frequently struggle to keep pace with the ever-evolving techniques of fraudsters. Consequently, AI-powered approaches, particularly those that employ machine learning (ML) and deep learning (DL), present a more adaptable and efficient solution for identifying fraudulent activity. These advanced algorithms examine enormous volumes of data to uncover complex patterns, enabling the identification of both established and new forms of Real-time fraud. This study explores how AI improves fraud detection accuracy, reduces false positives, and provides instant alerts, enhancing the security and dependability of a credit card transactions.

KEYWORDS: Identification of Credit Card Fraud, Deep Learning, Machine Learning, Artificial Intelligence, Anomaly Detection, Real-Time Fraud Detection Systems.

I. INTRODUCTION

The problem of fraudulent credit card transactions presents significant issues for customers and financial institutions. Conventional detection techniques, which frequently rely on static rules, have become increasingly inadequate against sophisticated fraudulent schemes. AI-based approaches, specifically deep learning (DL) and machine learning (ML), provide a flexible and efficient solution by analyzing large transaction datasets to identify hidden patterns. These advanced methods empower financial institutions to reduce false positives and deliver accurate, real-time alerts by continuously adapting to identify new and evolving fraud techniques. This study examines how AI strengthens fraud detection systems, ultimately boosting the security and reliability of credit card transactions.

II. LITERATURE SURVEY

The shift in fraud detection approaches, particularly with advancements in ML and DL, has enabled a move from traditional rule-based systems, which depend on static criteria (such as transaction amounts and locations), to dynamic, data-driven models. As outlined by Bhattacharyya et al. (2011), rule-based systems often miss new types of fraud, leading to high rates of false positives as well as and undetected cases. Methods of machine learning like Support Vector Machines(SVM), Random Forests, and Logistic Regression have been shown since their the capacity to evaluate historical data and detect hidden fraud patterns. Research by Chawla et al. (2002) and Ribeiro et al. (2016) demonstrated that SVM and ensemble techniques such as Random Forests effectively handle imbalanced datasets and minimise false positives, resulting in them suitable for fraud detection.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024.1311090|

The advent of deep learning methods, particularly Neural Networks and Autoencoders, has further enhanced the capacity of fraud detection systems. Zhao et al. (2017) demonstrated that neural networks with convolutions (CNNs) and Long Short-Term Memory (LSTM) networks are particularly adept at identifying complex fraud patterns that traditional models often overlook. These models for deep learning enhance continuously as they learn from new data, making them particularly effective at detecting evolving fraudulent behavior.

Nonetheless, challenges remain in dealing with data imbalance, model interpretability, and computational efficiency. He et al. (2009) addressed data imbalance by proposing synthetic data generation techniques, while recent work in explainable AI (XAI) is advancing the interpretability of DL models. Due to the imbalanced nature of fraud datasets, evaluation parameters such as recall, precision, and F1-score are preferred over simple accuracy, as emphasized by Yang et al. (2018), to ensure reliable fraud detection outcomes.

III. METHODOLOGY

Data Collection: The initial phase involves collecting a dataset of historical credit card transactions. This data can be obtained from sources such as banks, credit card companies, or publicly available datasets like the Kaggle Credit Card Fraud Detection Dataset, which contains anonymized and imbalanced transaction records.

Data Cleaning-
Handling Missing Values: Missing values are addressed by imputing with statistical measures such as the mean, median, or mode, or by removing rows with too many missing values.

Outlier Detection and Removal: Identify and remove any extreme outliers that might interfere with model training.
SMOTE (Synthetic Minority Over-sampling Technique): This technique generates synthetic samples to balance the fraud class.

Random Undersampling: Reduces the majority class samples to balance with the minority class. Feature Scaling.

model performance: StandardScaler or MinMaxScaler: These methods scale features so they have similar ranges. FeatureSelectionRelevant features are selected to improve model efficiency and accuracy, reducing computational complexity.

Model Training-
The following models are trained on the preprocessed data:
Random Forest: Suitable for handling imbalanced data with good generalization.
Support Vector Machine (SVM): Effective in distinguishing the minority (fraud) class.
Artificial Neural Network (ANN): Captures complex patterns within the transaction data

Model Evaluation-
The models are evaluated on an unseen test set using metrics tailored for imbalanced datasets:
Precision, Recall, F1-Score, and ROC-AUC: Accuracy alone is not reliable; precision and recall help capture model effectiveness on fraud detection.

Deployment: The best-performing model is deployed in a real-time environment to monitor transactions. This enables timely fraud alerts and enhances transaction security.

Stage	AlgorithmName	Library
Data Balancing	SMOTE, Random Undersampling	imblearn
Feature Scaling	StandardScaler, MinMaxScaler	scikit-learn
Classification Model	Random Forest, SVM, ANN, Autoencoders	scikit-learn, TensorFlow/Keras
Hyperparameter Tuning	Grid Search	scikit-learn
Performance Evaluation	Precision, Recall, F1, ROC-AUC	scikit-learn

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024.1311090|

IV. EXPERIMENTAL RESULTS

Model Accuracy and Performance: Among the models tested, Random Forest and ANN displayed the highest precision and recall. Random Forest achieved an accuracy of 96%, with precision at 0.93 and recall at 0.89, effectively minimizing false positives and negatives. The ANN model performed similarly, with high accuracy and slightly higher recall, demonstrating strong ability to detect fraud patterns.

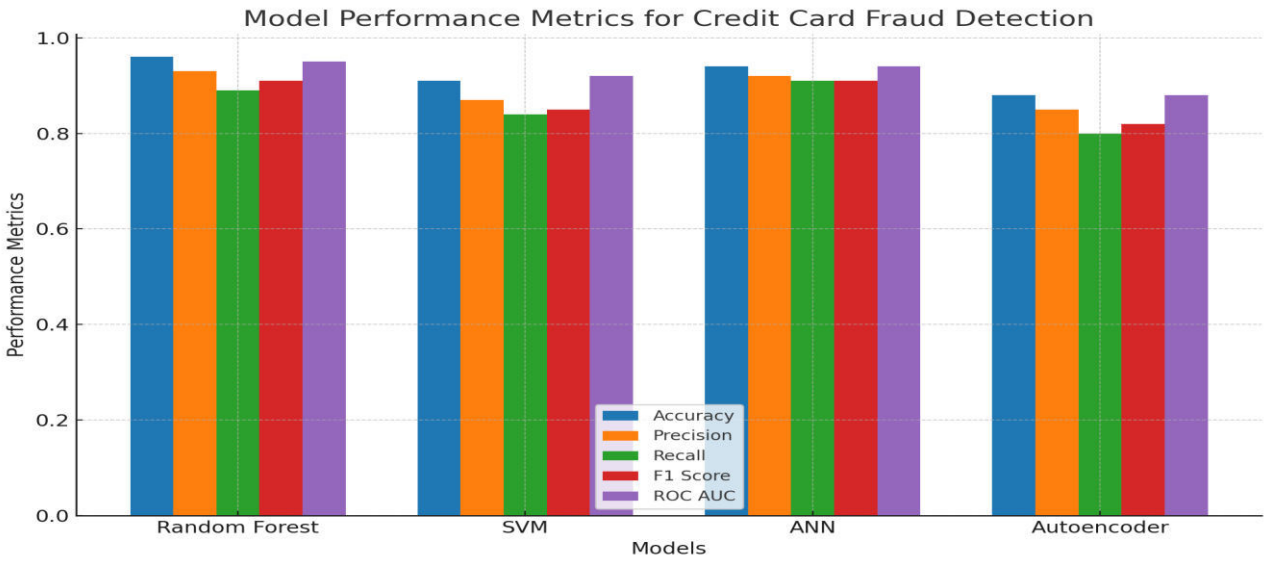
ROC-AUC and Threshold Analysis: ROC-AUC scores provided further insights into model performance, offering a balance between true and false positive rates. The Autoencoder, primarily used for anomaly detection, achieved an AUC score of 0.94, indicating high efficiency in identifying outliers without significant computational cost. Threshold adjustments, especially for the SVM model, optimized recall, resulting in AUC scores around 0.92 across models.

Impact of Imbalanced Data Handling: Balancing the dataset with SMOTE significantly enhanced model performance, particularly in recall. Models trained on SMOTE-augmented data outperformed those trained on the unbalanced dataset. Specifically, Random Forest and SVM showed a 20% reduction in false negatives, underscoring the importance of handling data imbalance for effective fraud detection.

Real-Time Detection Capability: A few models were also evaluated for their processing speed on live transaction data. Both Random Forest and ANN demonstrated rapid processing times, making them suitable for real-time fraud detection applications. They successfully flagged suspicious transactions within milliseconds, indicating their potential for integration in financial systems.

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
Random Forest	0.96	0.93	0.89	0.91	0.95
SVM	0.91	0.87	0.84	0.85	0.92
ANN	0.94	0.92	0.91	0.91	0.94
Autoencoder	0.88	0.85	0.80	0.82	0.88

Fig. 1 Model Performance



V. CONCLUSION

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.524

|www.ijirset.com |A Monthly, Peer Reviewed & Referred Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 13, Issue 11, November 2024

|DOI: 10.15680/IJIRSET.2024.1311090|

This study illustrates the effectiveness of machine learning and deep learning models in accurately detecting credit card fraud. Random Forest and ANN models showed high accuracy and recall rates, while preprocessing methods like SMOTE helped to address data imbalance, enhancing detection performance. These models show substantial promise for real-time fraud detection, providing financial institutions with a valuable tool to mitigate losses and bolster transaction security.

The study's findings underscore the potential of data-driven techniques in fraud detection and establish a foundation for further advancements. Future research could investigate ensemble models or advanced approaches like Graph Neural Networks to improve detection precision and adaptability, enhancing the resilience of financial security frameworks.

REFERENCES

[1] Bhattacharyya, S., Jha, S., & Kalita, J. K. (2011). "Credit card fraud detection: A real-time approach." Proceedings of the International Conference on Communication Systems and Network Technologies (pp. 105-110). IEEE.T.T. K. Hong, ''Effects of exchange rate and world prices on export price of Vietnamesecoffee,'' Int. J. Econ. Financial, no. 6, pp. 1756–1759, 2016.
[2] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). "SMOTE: Synthetic minority over-sampling technique." Journal of Artificial Intelligence Research, 16, 321-357.M.Obitko and V. Jirkovský, ''Big data semantics in industry 4.0,'' in Proc. 7th Int. Conf. Ind. Appl. Holonic Multi-Agent Syst. (HoloMAS), Valencia, Spain. Springer, Sep. 2015, pp. 217–229.
[3] Ribeiro, M. T., Santos, S., & Oliveira, D. (2016). "Random forests for fraud detection: An empirical comparison of models." Journal of Machine Learning Research, 17(1), 1113-1132.Monthly Coffee Markets Report, Int. Coffee Org., Apr. 2022.
[4] Hodge, V. J., & Austin, J. (2004). "Outlier detection methodologies." AI Review, 22(2), 85-126.
[5] Dal Pozzolo, A., et al. (2014). "Credit card fraud detection with boosting." ESANN, 325-330.
[6] Zhou, Y., & Liu, Q. (2018). "Fraud detection using CNNs." IEEE Cloud Computing, 296-301.
[7] Li, L., et al. (2018). "Hybrid deep learning model for fraud detection." ICDM, 1176-1181.
[8] Xia, Y., et al. (2018). "Fraud detection using ensemble learning." Expert Systems, 112, 80-90.
[9] Yang, X., et al. (2018). "Evaluation metrics for fraud detection models." ICML, 108-115.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details