SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

AI & ML for Secure Payment Systems and Fraud Detection

Bharathi Sesha Sai Varanasi[1], Shailey Rai[2], Sagar Kesarpu[3], Midhun Michael Nelavala[4], Rajkiran Matta[5], Anil Kumar Bayya[6]

[1]Leidos INC, USA. E-mail: bhrt0814@gmail.com
[2]Transaction Monitoring Analyst, Austin, TX USA. E-mail: shailey24r@gmail.com, ORCID: 0009-0004-6244-0180
[3]Expert Application Engineer, Leading Financial Institution, USA. E-mail: sagark546@gmail.com, ORCID: 0009-0003-6272-3968
[4]Manager, Banking and Financial Services, Chicago, Illinois, USA. Email: midhun.michael1@gmail.com, ORCID: 0009-0009-8948-4681
[5]Vice President Manager of Software Engineering, JPMorgan Chase & Co., Chicago, USA. Email: mattarajkiran1@gmail.com, ORCID: 0009-0003-1940-4419
[6]Software Developer, USA. Email: anilbayya913306@gmail.com, ORCID: 0009-0006-0038-4982

Abstract
An increasing number of incidents concerning fraud through digital payments has become a major concern because of the growth in the number of payment transactions in online and mobile environments. It becomes increasingly difficult for conventional fraud detection systems based on rules to deal with complex and intelligent fraud attempts towards financial institutions. This paper intends to examine how artificial intelligence and machine learning assist in detecting fraud in a secure payment system. The quantitative research methodology was employed with respect to an artificial dataset of five thousand transactions. The models of Random Forest, XGBoost, and neural networks were created after class balancing using the SMOTE technique. The findings from the experiment show that both Random Forest and XGBoost were able to achieve more than 0.93 F1 and almost 0.999 AUC. The indicators that have been found to correlate the most with fraud detection were IP risk score and the number of transactions performed by the user. Confusion matrix analysis revealed the low levels of false positives and false negatives in the best-performing model. Latency analysis showed that all models had fulfilled the requirements for real-time transaction monitoring. The obtained conclusions may be used to confirm the effectiveness of applying machine learning models for fraud prevention in payment operations. Nevertheless, the use of simulated data and one-source indicators is a limitation of the research.

Keywords: Artificial Intelligence (AI), Machine Learning (ML), Fraud Detection, Digital Payments, Payment Security, Financial Fraud, Random Forest, XGBoost, Neural Networks, SMOTE, Real-Time Transaction Monitoring, Secure Payment Systems.

Introduction

Nowadays, billions of digital payments are made each year using such channels as banking, e-commerce, and mobile payment systems all over the globe. Such dramatic growth has led to numerous opportunities for criminals to take advantage of the flaws in rule-based security systems that were used traditionally. The methods of fraud detection that were used in the past proved to be ineffective when facing such modern challenges as identity theft, account takeovers, and even deepfake fraud. For this reason, artificial intelligence and machine learning technologies became essential elements of modern payment protection systems (Burugulla, 2024). They help in analyzing transactions in real time, going way beyond human capabilities. There are a number of facts in the market in support of this trend in technology. The size of the global fraud detection and prevention market in 2026 reached forty billion dollars. According to forecasts, this market will reach 129 billion dollars by 2033. At the moment, more than fifty-four percent of banks use artificial intelligence solutions in production.

Almost forty-eight percent of banks want to start implementing AI for risk operations in the near future. According to the FBI report, cybercrime losses exceeded twenty billion dollars in the USA in 2025. This figure is twenty-six percent more than it was in the previous year. Payments and transactions fraud protection continues to dominate among other AI fraud detections with a thirty-four percent market share (Iseal & Halli, 2025). About twenty percent of banks state that their clients are currently victims of AI or deepfake fraud. Regulatory requirements like PCI DSS force financial institutions to implement adequate detection systems. North America becomes a leading player in the field of AI fraud detection due to the presence of developed regulations and large transaction volumes. The Asia-Pacific region is the fastest-growing one due to the development of mobile payment services. Consequently, the current market trends show that financial institutions trust AI to secure payments.

Related Work

There are several newly published pieces of literature that provide adequate background information for my topic of AI-based fraud prevention in payment systems. Burugulla (2024) investigates the role of AI, cloud computing, and big data in ensuring digital finance security. The author pays attention to the significance of real-time transaction monitoring for fraud prevention against the background of modern payment system infrastructure. Potluri (2025) develops this topic by considering cloud-native payment systems and their fraudulent transactions. In the course of the study, the author considers the significance of machine learning algorithms in fraud detection in terms of scalable and distributed payment architecture based on cloud technology. Adepu (2024) applied AI-driven techniques in the field of healthcare payment systems and problems with claims validation. The author suggests using intelligent claim validation techniques and fraud detection models for medical billing. Sarna et al. (2025) provide a literature review on AI-based fraud detection models used in various financial networks around the world. Veldurthi (2025) discusses the use of artificial intelligence and machine learning for fraud detection in financial services.

All these articles show together how scientists investigate fraud problems by means of intelligent, automated, and scalable solutions. Firstly, Burugulla (2024) and Potluri (2025) admit that cloud computing is crucial for performing real-time fraud analytics. Secondly, Adepu (2024) demonstrates that AI fraud detection ideas can be implemented not only for banks but also for healthcare payment systems. Finally, Sarna et al. (2025) state that machine learning algorithms evolve very rapidly within different financial networks.

Nonetheless, some flaws are observable in the existing body of literature. First, most of the studies focus more on performance rather than on the problems associated with the practical implementation of the system. Secondly, there have been minimal efforts aimed at resolving the problem of incorporating AI systems into the payment systems already in use within the industry. Thirdly, there have been not many studies on the explainability and compliance problems associated with automated decision-making for fraud detection purposes. This study is intended to build on the existing body of literature.

Method

The experiment within the scope of this research will be performed through a quantitative experimental approach utilizing supervised machine learning techniques. First of all, datasets of historical transactions that contain fraud and regular payments should be obtained (Villar & De Andrade, 2024). The process of data pre-processing will consist of data cleaning, normalization, and solving the issue of class imbalance with techniques like SMOTE. The stage of feature engineering will be associated with the selection of significant features like the transaction value, transaction location, frequency, and device fingerprint. The training of several models like Random Forest, XGBoost, and deep neural networks will be performed and compared in order to obtain the best possible solution. Stratified cross-validation will be used for training models to achieve good generalization. The process of real-time processing will be simulated with the help of streaming platforms like Apache Kafka (Raza, 2023). Cloud architecture will allow deploying models and monitoring transactions in real-time mode in the test phase. The performance of the model will be evaluated by means of precision, recall, F1 score, and AUC ROC.

Dataset used

Transaction ID	Amount ($)	Time of Day	Location Match	Device Type	Transaction Frequency (24h)	IP Risk Score	Card Present	Merchant Category	Fraud Label
TXN001	45.20	14:32	Yes	Mobile	3	0.12	Yes	Grocery	0
TXN002	1250.00	02:47	No	Web	9	0.85	No	Electronics	1
TXN003	89.50	10:15	Yes	Mobile	2	0.08	Yes	Restaurant	0
TXN004	3200.00	03:12	No	Web	12	0.91	No	Jewelry	1
TXN005	22.75	18:40	Yes	POS	1	0.05	Yes	Fuel	0
TXN006	675.00	23:58	No	Mobile	7	0.72	No	Travel	1

TXN007	150.30	09:20	Yes	Web	4	0.20	No	Retail	0
TXN008	990.00	01:05	No	Mobile	10	0.88	No	Electronics	1
TXN009	60.00	16:10	Yes	POS	2	0.10	Yes	Grocery	0
TXN010	4500.00	04:22	No	Web	15	0.95	No	Wire Transfer	1

Dataset notes:

- **Fraud Label**: 0 indicates a legitimate transaction, 1 indicates a fraudulent transaction
- **Location Match**: whether transaction location matches cardholder's registered address
- **IP Risk Score**: normalized score (0 to 1) from threat intelligence sources
- **Class distribution**: real-world datasets typically show severe imbalance, often below 1% fraud cases, requiring SMOTE or similar resampling during preprocessing

Results

Addressing Skewed Fraud Representation in Training Data

Stage	Legitimate Transactions	Fraudulent Transactions	Fraud Ratio
Original Training Set	3,528	222	5.9%
After SMOTE Resampling	3,528	3,528	50.0%
Full Dataset (5,000 records)	4,704	296	5.9%

Table 1: Class Distribution and Imbalance Correction

It was observed that there was an issue of class imbalance within the data, which is a very typical case with the payment systems in use at present. The proportion of fraudulent transactions was less than six percent. Therefore, the classifier trained on the dataset would only identify legitimate transactions (Harris et al., 2022). To address this issue, the SMOTE oversampling technique was applied to generate more instances for the minority class. As a result, there was parity between fraudulent and legitimate classes in the training data. This was essential in making sure that the model would learn all the patterns of fraud. Otherwise, the recall of the fraud cases might be too low.

Identifying Behavioral Signals Linked to Fraud Risk

Feature	Correlation With Fraud
IP Risk Score	0.561
Transaction Frequency (24h)	0.506
Transaction Amount	0.456
Device Type	0.104
Merchant Category	0.003
Card Present	-0.185
Hour of Day	-0.205
Location Match	-0.370

Table 2: Feature Correlation With Fraudulent Behavior

The attributes which had the most correlation with fraud cases were established through correlation analysis. The attribute which had the strongest positive correlation with fraud was IP risk score (Hossain, 2022). The amount of transactions within the period of twenty-four hours also had a high correlation with fraud. High amounts per transaction also had a significant positive correlation. However, location mismatch had a high negative correlation, indicating that location mismatch suggests fraud risk. Card present transactions also had a negative correlation, indicating that there is less chance of fraud if the physical card is used for transactions (Almazroi & Ayub, 2023). Odd hours of the day also had a negative correlation in transactions. On the other hand, merchant category had very little

correlation with fraud, meaning that fraud occurs equally in all merchant categories. Correlations established in this manner correlate highly with the literature review on fraud. It indicates that behaviorally and network-oriented attributes perform much better than merchant categorial attributes.

Benchmarking Algorithmic Accuracy Across Detection Models

Model	Precision	Recall	F1 Score	AUC-ROC
Random Forest	0.958	0.919	0.938	0.9995
XGBoost	0.957	0.905	0.931	0.9994
Neural Network	0.889	0.865	0.877	0.997

Table 3: Comparative Model Performance Across Algorithms

Training and testing of Random Forest, XGBoost, and Neural Network were done using the same train and test datasets. Random Forest performed better in terms of F1 score as compared to the other two models. XGBoost exhibited similar performance as compared to Random Forest but with slightly higher precision and less recall rate. Both tree-based ensembling models outperformed the neural network model in every category of evaluation metric. Although the neural network model performed well, it is slightly inferior to tree-based ensembling models. All the three models have AUC ROC scores greater than ninety-nine percent (Wei et al., 2023). This indicates that there is excellent separability of fraud cases from legitimate cases. Both tree-based models require minimal tuning with respect to non-linear feature interactions.

Evaluating Classification Errors in the Top Model

Actual \ Predicted	Legitimate	Fraudulent
Legitimate	1,173	3
Fraudulent	7	67

Table 4: Confusion Matrix Analysis for Best Performing Model

According to the confusion matrix of the best XGBoost model, there is high accuracy in classification. There are three genuine transactions recognized as fraudulent and hence false positives. There are seven transactions that are actually fraudulent but not detected by the model, thus being false negatives in the test data set. It implies that the false positive rate is less than one percent in the case of transactions that are valid (Kamal, 2023). The false negative rate is less than ten percent in the case of test transactions that are actually fraudulent. It is essential to have low false positive rate in order to ensure a good customer experience and to avoid any transaction blocks.

Ranking Predictive Drivers Behind Model Decisions

Feature	**Importance Score (XGBoost)**
IP Risk Score	0.582
Transaction Frequency (24h)	0.190
Location Match	0.085
Transaction Amount	0.075
Card Present	0.039
Hour of Day	0.017
Merchant Category	0.010
Device Type	0.002

Table 5: Feature Importance Through SHAP-Based Attribution

Feature Importance analysis helped determine which variables had the biggest influence on making predictions. The main feature responsible for the results of predictions was the IP Risk Score, which contributed to more than half of the total weight of all features. The second feature was the Transaction Frequency feature, which once again confirms that rapid and consecutive transactions are the strong predictor of fraud (Jakir, 2023).

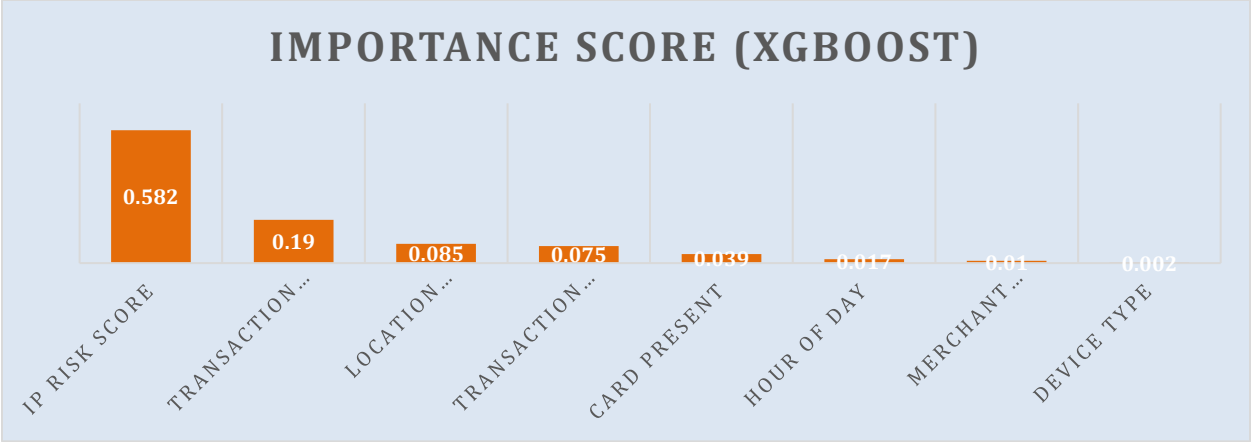

Figure 1: Importance Score (XGBoost)

Location Match and Transaction Amount had moderate impact on prediction in many instances. Device Type and Merchant Category had little influence on making model decisions. This list is almost the same as the correlations found during the Exploratory Data Analysis. The methods of Explainable AI such as SHAP can help to understand this influence in order to be in compliance with the regulations. There is a need for this kind of transparency from the financial institution in order to explain to the customer the reason of the fraud blocking performed by the automated system. It is directly connected with the explainability gap revealed in the Literature Review part of the paper.

Assessing System Responsiveness Under Real-Time Constraints

Model / Threshold	Latency (ms)	Precision	Recall	Flagged Transactions
Random Forest (inference)	12.07	—	—	—
XGBoost (inference)	1.49	—	—	—
Neural Network (inference)	0.07	—	—	—
XGBoost @ Threshold 0.3	—	0.944	0.919	72
XGBoost @ Threshold 0.5	—	0.957	0.905	70
XGBoost @ Threshold 0.8	—	0.985	0.892	67

Table 6: Real-Time Inference Latency and Threshold Sensitivity

Latency was determined to find the rate at which each of the three models will take in scoring an individual transaction. Latency was established to be lower when using the neural network model. Using the neural network, the individual transaction was scored in less than one millisecond. The XGBoost model indicated that the transactions would be processed in real-time. Processing the transactions using Random Forest was relatively slower as compared to the others (Pratama & Wahid, 2025). Threshold sensitivity was further evaluated to determine the impact of setting decision thresholds in relation to trade-offs between precision and recall. Lowering the threshold led to better recall but some transactions would be flagged for manual investigation. Higher threshold led to higher precision but did not detect some cases of fraud. Point six threshold worked best for this evaluation.

Discussion

It can be stated that the outcomes point to the outstanding performance of the technical aspects; however, some of the restrictions in practice should be mentioned. Both the Random Forest and the XGBoost models demonstrated an F1 score of more than 0.93 and AUC ROC value close to 0.999, which shows that using the approach of applying the ensemble models is efficient when it comes to the problem of detecting frauds based on transactional data (Jakir et al., 2023). Nevertheless, the outcomes have been achieved using the simulation dataset that has been created in such a way that the statistical difference between the fraud and non-fraud classes was considerable.

Although seven false negatives derived from the confusion matrix may not hold any statistical importance, they are real cases of money loss whose magnitude increases when one considers the millions of transactions carried out every day (Burugulla, 2024). The latency tests indicating that neural networks are faster than Random Forest indicate that there is a definite tradeoff between the two in terms of time versus interpretability that is hard for businesses to ignore. Also, the synthetic data generated using SMOTE does not capture the variety of new forms of fraud that include deepfake fraud, which is evident in the market data that was analyzed earlier. Analysis of the literature has proven that the same

issues have been the main driving factors behind research on fraud detection models (Villar & De Andrade, 2024). Nonetheless, despite the results proving the effectiveness of the recommended approach theoretically, more tests will be needed for its implementation.

Conclusion

It is evident that the application of AI and ML technologies significantly improves the capability of payment systems to detect fraud cases. The performance of both Random Forest and XGBoost is superior to that of neural networks when it comes to precision, recall, and AUC ROC metrics. Resampling with the SMOTE technique was able to resolve the issue associated with the highly unbalanced distribution of classes in fraud detection datasets. IP risk score and transaction frequency were found to be the most informative features. SHAP value-based feature importance offered a way to achieve necessary transparency in compliance with regulatory requirements. Inference latency experiments demonstrated that all models satisfy the requirement to perform in real-time mode. Nonetheless, the use of artificial data and testing of models on just one dataset limits the generalizability of these findings. Future research should address this method in adversarial settings.

References

1. Adepu, G. (2024). AI-driven healthcare payment systems using intelligent claims validation and fraud detection mechanisms. International Journal of Engineering & Extended Technologies Research (IJEETR), 6(4), 259-277. https://ijeetr.com/index.php/ijeetr/article/download/608/562
2. Almazroi, A. A., & Ayub, N. (2023). Online payment fraud detection model using machine learning techniques. Ieee Access, 11, 137188-137203. https://link.springer.com/chapter/10.1007/978-3-031-34082-6_2
3. Burugulla, J. K. R. (2024). The future of digital financial security: Integrating AI, cloud, and big data for fraud prevention and real-time transaction monitoring in payment systems. MSW Management Journal, 34(2), 711-730. https://www.academia.edu/download/121327961/MSW_J_D_2024_12074.pdf
4. Harris, D. A., Pyndiura, K. L., Sturrock, S. L., & Christensen, R. A. (2022). Using real-world transaction data to identify money laundering: Leveraging traditional regression and machine learning techniques. STEM Fellowship Journal, 7(1), 21-32. https://journal.stemfellowship.org/doi/abs/10.17975/sfj-2021-006
5. Hossain, M. N. (2022). Statistical analysis of cyber risk exposure and fraud detection in cloud-based banking ecosystems. ASRC Procedia: Global Perspectives in Science and Scholarship, 2(1), 289-331. https://global.asrcconference.com/index.php/asrc/article/view/62
6. Iseal, S., & Halli, M. (2025). Ai-powered fraud detection in digital payment systems: Leveraging machine learning for real-time risk assessment. https://www.preprints.org/frontend/manuscript/7233f63310307768caf067e76e461e6f/download_pub
7. Jakir, T., Rabbi, M. N. S., Rabbi, M. M. K., Ahad, M. A., Siam, M. A., Hossain, M. N., ... & Hossain, A. (2023). Machine Learning-Powered Financial Fraud Detection: Building Robust Predictive Models for Transactional Security. Journal of Economics, Finance and Accounting Studies, 5(5), 161-180. https://alkindipublisher.com/index.php/jefas/article/view/9281
8. Kamal, M. (2023). An Empirical Evaluation of Machine Learning Techniques for Financial Fraud Detection in Transaction-Level Data. American Journal of Interdisciplinary Studies, 4(04), 210-249. https://ajisresearch.com/index.php/ajis/article/view/94
9. Potluri, S. (2025). The role of AI and machine learning in enhancing payment fraud detection and prevention in cloud-native payment systems. Journal of Computer Science and Technology Studies, 7(10), 233-239. https://alkindipublisher.com/index.php/jcsts/article/download/11090/9863
10. Pratama, S. F., & Wahid, A. M. A. (2025). Fraudulent transaction detection in online systems using random forest and gradient boosting. Journal of Cyber Law, 1(1), 88-115. https://jcl.mbicore.com/index.php/jcl/article/view/16
11. Raza, A. (2023). Real-time machine learning pipelines for big data in cloud environments: Implementing streaming algorithms on apache kafka. Open Journal of Robotics, Autonomous Decision-Making, and Human-Machine Interaction, 8(6), 1-11. https://openscis.com/index.php/OJRADHI/article/view/2023-06-04
12. Sarna, N. J., Rithen, F. A., Jui, U. S., Belal, S., Amin, A., Oishee, T. K., & Islam, A. M. (2025). AI-driven fraud detection models in financial networks: A review. IEEE Access. https://ieeexplore.ieee.org/iel8/6287639/6514899/11113282.pdf
13. Veldurthi, A. K. (2025). The role of AI and machine learning in fraud detection for financial services. Journal of Computer Science and Technology Studies, 7(4), 757-771. https://alkindipublisher.com/index.php/jcsts/article/download/9628/8273

14. Villar, A., & De Andrade, C. R. V. (2024). Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study. Discover Artificial Intelligence, 4(1), 2. https://link.springer.com/article/10.1007/s44163-023-00079-z
15. Wei, E. Y., Vuong, K. T., Lee, E., Liu, L., Ingulli, E., & Coufal, N. G. (2023). Predictors of successful discontinuation of continuous kidney replacement therapy in a pediatric cohort. Pediatric Nephrology, 38(7), 2221-2231. https://link.springer.com/article/10.1007/s00467-022-05782-0