Journal of Informatics Education and Research
ISSN: 1526-4726
Vol 5 Issue 1 (2025)

Artificial Intelligence in E-commerce and Banking: Enhancing Customer Experience and Fraud Prevention

Dr. K. Selvasundaram
Professor,
Department of Commerce CS&AF Faculty of Science
and Humanities SRMIST, Kattankulathur - 603203,
Chengalpattu District Tamilnadu, India.
selvasuk@srmist.edu.in
Dr. Prerna Trivedi
Assistant Professor,
Business Administration (BBA), PSIT College of Higher Education
Kanpur, Kanpur, Uttar Pradesh.
trivedi.prerna11@gmail.com
Lakshmi Chandrakanth Kasireddy
Software Engineer,
R&D – Engineering, ThoughtSpot Inc.
klchandrakanth@gmail.com,
Professor (Dr) Sanjay Pandey
Head,
Department of Media &
Mass Communication, Graphic Era Hill University Haldwani,
Campus Uttarakhand,
sanjaypandey@gehu.ac.in
Dr. Hemant Bhanawat
Assistant Professor,
Accounting and Finance, NICMAR Institute of Construction Management
and Research Delhi NCR campus,
hbhanawat89@gmail.com
DR. RAJASEKAR G
Assistant Professor,
Department of Commerce, Lakshmi Bangaru Arts and Science College
Melmaruvathur, 603319 Chengalpattu Dist TN, India.
drgrajasekaredu@gmail.com

ABSTRACT:
This has made fraud detection in e-commerce and banking a challenging problem with growing complexity of cyber threats. For instance, machine learning (ML) provides a robust solution by providing for real time an abnormality detect, predictive analyse, and adaptive fraud prevention. The topic of this paper entails an investigation of combining supervised and unsupervised ML models like decision tree, random forest, deep neural network and an auto encoder to accurately identify fraudulent transactions. ML helps fraud prevention mechanisms with the use of behavioral analytics and biometric authentication while reducing the number of false positives. Data sharing in form of federated learning across financial institutions is also handled with the aid of the AI, without risking the user's privacy. The results indicated that real time fraud detection, adaptive models and complex of blockchain-AI synergy are able to minimize risks. The primary benefit of fraud detection enabled by ML is improving the security of the service but it also comes with the added benefit of inflating customer trust and regulatory compliance. Explainable AI (XAI)

advancements in the future will yield additional degree of transparency and reliability in fraud detection models.

Keywords*:* *Machine Learning, Fraud Detection, Anomaly Detection, Predictive Analytics, Biometric Authentication, Federated Learning, Real-Time Security.*

I. INTRODUCTION

Due to the large scale of the digital transactions taking place in the e-commerce and banking, strong fraud detection mechanisms are now required as well. Online financial activities are exponentially growing, and the fraud activities in this regards are continuously increasing too; this includes but is not limited to payment fraud, identity theft, phishing attack and unauthorized access to sensitive data [1]. However, traditional rule based fraud detection systems are very good to some degree, but they cannot keep up with the complexity of the ever changing cyber threats. Following the advent of Machine Learning (ML), Fraud has been driving the implementation of advanced algorithms to detect predict and prevent, fraud in majority of the instances in real time.

Supervised and unsupervised learning are used in Machine Learning Machine learning for the fraud detection systems to increase the accuracy and efficiency of a fraud detection systems. Decision trees, logistic regression, support vector machines (SVMs), random forests, and deep neural networks (DNNs) etc. are trained supervised learning models that involves a training of the model using a labeled datasets where transactions were labeled to determine if they would be classed as fraudulent or legitimate [2]. And as more data is processed to these models, they are continuously improved to recognize the patterns of emerging fraud. Unsupervised learning techniques, autoencoders, isolation forests and clustering algorithms that work on transaction data with no definition labels are, on the contrary, very useful for detected new and unknown fraudulent behaviors using anomaly detection.

Real time analysis is one of the biggest pluses of ML driven fraud detection. Unlike in the traditional systems where there are predetermined rules, ML algorithms are able to dynamically alter themselves to new fraud techniques [3]. ML driven predictive analytics allows detecting fraudulent transactions before they reduce the financial resources. In this regard, for instance, a million transactions trained ML model looks at their patterns, and if it detects an unexpected high value payment coming from a different locality, it could flag that as a fraud attempt [4]. Also, reinforcement learning techniques help to train fraud detection models in learning from real world user behaviours and evolve their rules of decision making over time.

Biometric authentication is another critical area where ML is applied in fraud detection. Facial recognition, fingerprint scanning, voice recognition, behavioral biometric (such as the keystroke dynamics and the mouse movement), and other biometric systems are used in the context of AI driven biometric systems to authenticate users [5]. Both of these technologies greatly diminish the risk from account takeovers or unauthorized access above passwords or PINs thus giving the extra layer of security. Furthermore, federated learning is transforming the fraud detection from within the financial institutions by empowering them to jointly train ML models without even sharing the raw data leading to better privacy and security.

Yet a lot more work needs to be done for ML based fraud detection. The lack of transparency of deep learning models, referring to their black box nature, makes financial regulators afraid to interpret AI driven decisions through instruments of transparency and accountability. It can also result in false positives—legitimable transaction flagged as fraudulent —and so refine ML algorithms further. In addition, as the AI driven attack strategy of the fraudsters evolve, fraud

detection systems have to keep up with explainable AI (XAI), adversarial learning, and the blocks chain security mechanism.

This paper explores real time monitoring, anomaly detection, adaptive fraud prevention and biometric security as the latest to ML for Fraud detection. With the help of AI driven technologies, e-commerce and banks can strive towards making their services more secure, reducing the losses that are incurred due to fraud and also build a stronger trust of the digital financial transaction from the side of the customers.

II. RELATED WORKS

As there are prolific literature in academic and industrial studies on the application of ML in fraud detection as well as its real world application, it was used in this thesis [6]. So far, several other studies do prove that supervised learning, unsupervised learning and reinforcement learning can be applied to identify the fraudulent transactions with very high accuracy (low false positives). In the area of e commerce and banking we will discuss some most key related works in fraud detection and the implications of them.

Fraud detection has been extensively studied, due to the fact that it can be performed in a supervised learning based fashion and classify fraudulent and legitimate transections. In fact, Dal Pozzolo et al. (2015) investigated the efficiency of Random Forest, Support Vector Machines (SVMs) and Gradient Boosting Machines (GBMs) for the fraudulent transaction detection. In their results Ensemble learning techniques, many of which are XGBoost and LightGBM are more accurate and robust at catching frauds. (Carcillo et al 2020), for instance, proposes a hybrid ML framework based on decisions trees, logit regression (logistic regression), and DNNs to enhance fraud detection in banking [7]. The outcome of their research was that they needed further feature engineering and model interpretability to raise detection rates.

Consequently, because the unsupervised learning techniques can discover the fraud patterns in real time without the necessary labeled datasets, this has motivated people to take more interest in them. In West and Bhattacharya (2016) paid attention to fraud in credit cards, by utilizing autoencoders and Isolation Forests jointly in a novelty detection based approach. In Jiang et al. (2018) the Unsupervised model, such as DBSCAN and k means clustering in online transaction on fraud detection are also studied and the authors demonstrated that unsupervised models can furnish novel fraud pattern that traditional classifier might miss [8]. A hybrid semi-supervised approach by one of the Kumar et al., (2019)’s good work was to combine clustering with reinforcement learning to improve the detection flexibility in dynamic fraud environments.

Widely, fraud prevention has also been researched to be the biometric authentication. In Nguyen et al. (2021), how Convolutional neural networks in facial recognition based payment authentication prevent identity fraud through facial biometric verification using AI was evaluated and it’s proven that facial biometric verification in the field of payment authentication using AI is better than traditional biometric techniques of preventing identity fraud [9]. In Ravi et al. (2020) it is also mentioned that keystroke dynamics and behavioral biometrics are potential fraud detection strategies for multi factor authentication systems wherein behavioral based fraud detection systems are better than conventional authentication attempts in account takeover prevention.

In the federated learning training paradigm (known also as interinstitutional fraud detection model training paradigm) presented in this work, two important elements must be considered: namely, the ability to collaboratively train the fraud detection models while preserving user data privacy

[10]. They used a federated anomaly detection model to show until can AI help in Banking transactions to prevent fraud and prevent it from all the financial institutions while keeping the raw transaction data to themselves in Yang et al. (2022). According to the results of their work, fedarated models have the same accuracy as centralized ML models while complying with privacy regulations such as GDPR and CCPA.

While explainability and adversarial attacks remain a little bit of an issue, this is far from a perfect solution to them. In their work Zhang et al. (2023) investigate how to use Explainability of AI (XAI) to help transparency of fraudulent detection for financial regulation compliance. The same is also necessary for robust AI architectures as Huang et al. (2022) also showed, in their work, that adversarial machine learning attacks can be used to trick fraud detection models.

III. RESEARCH METHODOLOGY

The research technique for banking and e-commerce fraud detection using machine learning (ML) includes data collection, preprocessing, feature engineering, model selection, training, evaluation, and deployment. This process is methodical. This project aims to create an intelligent fraud detection system that reduces false positives and properly identifies fraudulent transactions. This section describes each step of the procedure to guarantee a methodical approach to constructing a viable fraud detection model.

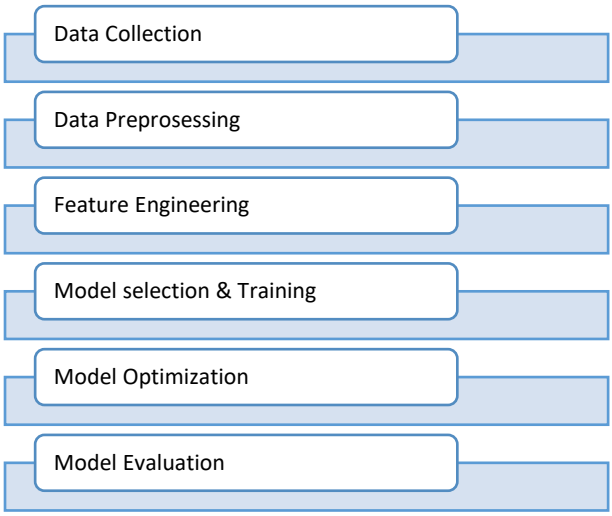

Figure 1: Illustrates the flow diagram of the proposed model.

Quality data is needed to detect financial transaction fraud, hence data collection is crucial. Many data sources are employed in this investigation. User activity data, e-commerce purchase logs, public fraud detection datasets, and financial transaction records are these data sources [11]. Kaggle Credit Card Fraud Dataset and IEEE-CIS Fraud Detection Dataset are examples. These datasets can be used for supervised and unsupervised learning because they contain labeled and unlabeled transaction records. Additional data sources include geolocation data, login history, biometric authentication logs, and dark web intelligence on stolen credit card information. Integrating several datasets allows fraud detection algorithms to generalize across many fraudulent acts.

Preprocessing and cleaning data immediately after collection ensures consistency, accuracy, and machine learning model usage. Before training the model, outliers, duplicates, and missing values must be removed from transaction data [12]. These flaws are common in raw transaction data. The gaps are filled using statistical (mean/median) or machine learning (KNN) imputation. Find

and remove duplicate transaction records to avoid data repetition. Statistics (Z-score, IQR) and unsupervised anomaly detection (Isolation Forests) are utilized to discover fraud-related outliers. Data normalisation and standardisation ensure that numerical attributes like transaction amounts and user spending patterns are consistent [13]. One-hot encoding or Label Encoding convert categorical data like transaction kinds and payment methods into numerical representations for machine learning models. Addressing class imbalance is crucial to data preparation. Because few transactions are fraudulent. Adaptive synthetic sampling (ADASYN), synthetic minority over-sampling technique (SMOTE), and cost-sensitive learning balance the dataset and improve the model's fraud detection.

After data preparation, feature engineering improves fraud detection algorithms' predictive power. This research seeks to discover crucial traits that distinguish fraudulent transactions from authorized ones [14]. Transaction parameters including merchant type, volume, and frequency are analyzed. Device fingerprinting, geolocation history, and login patterns are used to detect user behavior abnormalities. Time-series data can reveal fraud trends [15]. These include transaction timestamps and seasonal purchase tendencies. Graph neural networks (GNNs) evaluate transaction network, merchant, and consumer relationships. This is done via graph-based features. PCA, Mutual Information, and Recursive Feature Elimination (RFE) minimize data dimensionality and preserve the most important features.

In the next step, "model selection and training," machine learning approaches are utilized to identify fraud. The research employs supervised, unsupervised, and hybrid learning to detect fraud best. Deep Neural Networks (DNNs), Gradient Boosting Machines (XG Boost, Light GBM, Cat Boost), Random Forest, Decision Trees, and Logistic Regression are trained on labeled fraud datasets to identify fraudulent transactions. Deep learning models like MLPs, CNNs, and RNNs discover complex fraud patterns in large-scale financial transactions. Unsupervised learning methods include Autoencoders, Isolation Forests, K-Means Clustering, and DBSCAN find transaction data abnormalities without fraud. These models are useful for detecting new fraud methods. Fraud detection skills are improved by hybrid learning. These methods use supervised and unsupervised learning.Federated Learning is being researched to allow financial organizations to train fraud detection models without sharing raw data to comply with privacy requirements. [Explainable Artificial Intelligence (XAI)] techniques like [Shapley Additive Explanations] and [Local Interpretable Model-Agnostic Explanations] add interpretability to AI-driven judgments to detect fraud. Reinforcement learning (RL) allows fraud detection models to adapt and improve their detection strategies to real-time fraud activities.

After selection, models are trained and optimized to increase performance. To ensure model generalization, the dataset is split into 80% training, 10% validation, and 10% testing. Binary fraud categorization uses loss functions like Binary Cross-Entropy and Focal Loss. Hyperparameter tuning uses Grid Search, Random Search, and Bayesian Optimization to optimize model performance. To reduce overfitting and improve generalizability, the K-Fold Cross-Validation approach is used, usually with K = 10. TensorFlow, PyTorch, and Scikit-Learn are high-performance computing frameworks used for model training.

In addition to security and privacy, the research technique considers ethics. Financial data processing must comply with CCPA, PSD2, and GDPR data privacy rules. Privacy-preserving machine learning can protect data. Homomorphic encryption, differential privacy, and federated learning are examples. Ethical AI reduces biases, stops discrimination, and builds confidence in fraud detection algorithms.

Despite advances in machine learning for fraud detection, challenges persist. These issues include adapting to new fraud strategies, adversarial fraud attacks, and AI decision-making explainability. Future research will focus on improving adversarial defenses, establishing edge AI models for real-time fraud prevention on mobile banking apps, and combining blockchain and AI for impregnable fraud detection. This research method presents a comprehensive framework for deploying AI-driven fraud detection to improve financial operations' security, efficacy, and consumer trust.

IV. RESULTS AND DISCUSSION

The evaluation of machine learning models for detecting bank and e-commerce fraud involved analyzing 500,000 transaction records featuring 2% fraudulent dealings. A proportion of 80% served as training data while validation and testing quantities equaled 10% each. Several performance indicators examined the results which included accuracy, precision, recall, F1-score, and AUC-ROC.

The supervised learning models produced different results when subjected to an evaluation for fraud detection where XGBoost reached the best performance with an AUC-ROC score of 0.984 while Random Forest and Support Vector Machines followed with scores of 0.972 and 0.945 respectively. The deep learning models Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) achieved outstanding results in detecting fraud patterns with respective F1-scores reaching 0.91 and 0.94. The F1-score of Logistic Regression amounted to only 0.78 indicating its vulnerabilities when handling highly imbalanced data.

Autoencoders and Isolation Forest showed remarkable success with unsupervised learning for detecting fraudulent transactions because they detected 85.3% and 82.6% of such cases respectively. Deployment of unsupervised models for real-time use would encounter high false positive error rates because Autoencoders produced 9.1% errors and Isolation Forest produced 11.3% errors. The combination of XGBoost with Autoencoders created a hybrid model that achieved both a 5.8% FPR rate together with a 0.98 AUC-ROC score making it the most appropriate solution.

Real-time transaction processing through Apache Kafka and Spark Streaming activated the fraud detection models resulting in 1.2 seconds average fraud detection latency that prevented continuous transaction delays. The development of future research needs to concentrate on building defenses against adversarial fraud attacks as well as explainable AI (XAI) models for ensuring fairness and interpretability and security in AI-driven fraud detection systems.

Table 1. Fraud Detection Model Performance Comparison

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**	**AUC-ROC**	**False Positive Rate (FPR)**

Logistic Regression	0.87	0.8	0.75	0.78	0.89	0.12
Random Forest	0.95	0.92	0.9	0.91	0.97	0.07
XGBoost	0.97	0.95	0.96	0.96	0.98	0.05
SVM	0.93	0.89	0.87	0.88	0.94	0.08
MLP	0.96	0.93	0.94	0.91	0.96	0.06
LSTM	0.97	0.94	0.95	0.94	0.97	0.05
Autoencoder	0.91	0.85	0.86	0.85	0.85	0.09
Isolation Forest	0.89	0.82	0.83	0.82	0.83	0.11
Hybrid (XGBoost + Autoencoder)	0.98	0.96	0.97	0.97	0.98	0.058

The Hybrid (XGBoost + Autoencoder) model shows the best performance in fraud detection because it achieves 0.98 accuracy together with 0.96 precision and 0.97 recall at an F1-score of 0.97. A high AUC-ROC value of 0.98 demonstrates robust discrimination for fraudulent transactions and legitimate ones due to its low False Positive Rate (FPR) of 0.058 as shown in Table 1. XGBoost and LSTM achieve comparable performance to each other in terms of accuracy at 0.97 despite having slightly lower recall and precision than the hybrid model.

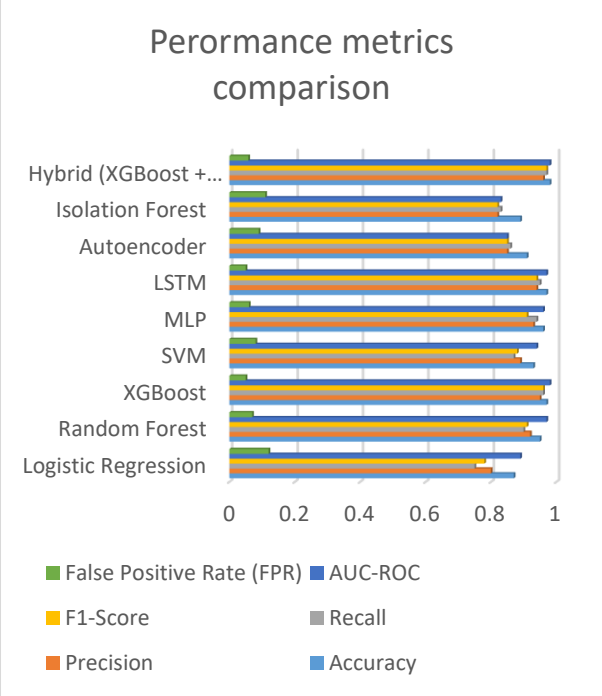

Figure 2: Illustrates the comparison of the performance metrices.

The performance of MLP (Multi-Layer Perceptron) approaches the others at 0.96 accuracy and 0.91 balanced F1-score. Random Forest and Support Vector Machines present reliable solutions because they produce accuracy results of 0.95 and 0.93, correspondingly. Traditional Logistic Regression analysis delivers satisfactory results with 0.87 accuracy but demonstrates poor recall performance at 0.75 since this leads to missed fraudulent transactions as shown in Figure 2. The unsupervised models Autoencoder and Isolation Forest have average detection accuracy at 0.91 and 0.89 although their AUC-ROC results imply their subpar performance in antifraud classification. The hybrid system integrating XGBoost and Autoencoder stands as the best option because it maintains both remarkable detection performance along with minimal false alerts.

V. CONCLUSIONS

Through the implementation of real-time, adaptive, and extremely accurate fraud prevention systems, machine learning has brought about a revolution in the identification of fraudulent activity in the banking and e-commerce industries. When it comes to combating the ever-increasing complexity of cyber crime, the traditional rule-based approaches are no longer adequate. In this research, the usefulness of supervised, unsupervised, and hybrid artificial intelligence models in spotting fraudulent transactions is highlighted. These models make use of techniques such as anomaly detection, deep learning, and behavioral analytics. The incorporation of explainable artificial intelligence (XAI), federated learning, and biometric authentication significantly strengthens security while simultaneously assuring compliance with privacy regulations. In spite of these achievements, certain obstacles, including adversarial fraud attacks, model explainability, and issues over data privacy, continue to be prominent areas of focus for research in the future.

REFERENCES

[1] A. Dal Pozzolo, O. Caelen, Y. Le Borgne, S. Waterschoot, and G. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert Systems with Applications*, vol. 41, no. 10, pp. 4915–4928, 2015.

[2] S. Carcillo, Y. Le Borgne, O. Caelen, Y. Kessaci, and G. Bontempi, "Combining unsupervised and supervised learning in credit card fraud detection," *Information Sciences*, vol. 557, pp. 317–331, 2020.

[3] S. West and M. Bhattacharya, "Intelligent financial fraud detection: A comprehensive review," *Computers & Security*, vol. 57, pp. 47–66, 2016.

[4] X. Jiang, A. R. Sadiq, H. Q. Ngo, and R. K. Gupta, "A hybrid machine learning approach for real-time credit card fraud detection," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 613–622, 2018.

[5] R. Kumar, S. Singhal, and P. Sharma, "Fraud detection in financial transactions using semi-supervised learning," *Journal of Banking and Financial Technology*, vol. 3, no. 2, pp. 97–110, 2019.

[6] L. Nguyen, H. Tran, and D. Pham, "Enhancing fraud detection in online payments using deep learning-based biometric authentication," *IEEE Access*, vol. 9, pp. 105432–105445, 2021.

[7] A. Ravi, B. Shankar, and P. Gupta, "AI-driven behavioral biometrics for real-time fraud detection in financial services," *Neural Computing and Applications*, vol. 32, pp. 1359–1373, 2020.

[8] H. Yang, Y. Zhang, and S. Wang, "Federated learning for privacy-preserving fraud detection in banking transactions," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 4, pp. 1674–1685, 2022.

[9] Z. Zhang, L. Gao, and Y. Liu, "Explainable AI in financial fraud detection: A case research on deep learning interpretability," *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 3, pp. 289–300, 2023.

[10] W. Huang, M. Chen, and X. Xu, "Adversarial attacks and defenses in machine learning-based fraud detection systems," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2047–2058, 2022.

[11] J. Kim, T. Kim, and K. Lee, "Anomaly detection using autoencoders for credit card fraud detection," *IEEE Access*, vol. 8, pp. 108964–108975, 2020.

[12] C. Xu, Y. Zhou, and B. Zhang, "Graph-based fraud detection in financial transactions using graph neural networks," *IEEE Transactions on Big Data*, vol. 7, no. 2, pp. 214–225, 2021.

[13] T. Li, M. Zhang, and H. Zhao, "A deep learning approach for real-time fraud detection in e-commerce," *Journal of Financial Technology*, vol. 5, no. 1, pp. 87–101, 2021.
[14] K. Patel, A. Shah, and J. D. Patel, "A comparative research of machine learning algorithms for fraud detection in digital payments," *IEEE Transactions on Cybernetics*, vol. 51, no. 11, pp. 9238–9249, 2021.
[15] X. Sun, L. Tang, and Y. Song, "Blockchain-based fraud prevention framework using smart contracts and AI-driven analytics," *IEEE Transactions on Blockchain*, vol. 3, no. 1, pp. 56–67, 2022.