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

Hybrid Blockchain and Machine Learning Framework for Real-Time Fraud Detection in Digital Transactions

Atharva Joshi, Om Pansare, Pranav Sarode, Tanish Jain, Rohan Shelke

Department of Computer Science and Engineering (Artificial Intelligence and Data Science)
Sanjivani University, India

Emails: {atharva.joshi24, om.pansare24, pranav.sarode24, tanish.jain24, rohan.shelke24} @sanjivani.edu.in

Abstract

Fraudulent activities in digital transactions have surged with the expansion of online banking, e-commerce, and financial services. Traditional fraud detection mechanisms are limited by centralization, susceptibility to data tampering, and delayed response times. This project aims to develop a secure, transparent, and intelligent fraud detection system by integrating Machine Learning (ML) and Blockchain technology. The system uses ML algorithms to detect anomalous or suspicious patterns in transactional data in real-time. Once analyzed, every transaction — along with its fraud status — is securely logged onto a Blockchain ledger, ensuring immutability and auditability of data. Additionally, smart contracts can automate actions like flagging accounts or sending alerts when fraud is detected. This hybrid approach enhances fraud prevention through predictive analytics while ensuring data integrity and transparency, making it suitable for use in banking, insurance, e-commerce, and government identity systems.

Introduction

In today's digital economy, fraud has become a growing threat across various sectors such as finance, healthcare, e-commerce, and insurance. The increasing volume and complexity of online transactions have created lucrative opportunities for malicious actors. Traditional fraud detection systems (FDS), often reliant on static rules and centralized databases, struggle to keep pace with dynamic and sophisticated fraudulent schemes. These legacy systems are plagued by several limitations: they are susceptible to single points of failure, vulnerable to internal data tampering, and often generate a high number of false positives, leading to poor customer experience. Furthermore, their reactive nature means fraud is often discovered long after it has occurred, resulting in significant financial losses.

The need for a paradigm shift is evident. There is a pressing demand for solutions that are not only intelligent and proactive but also inherently secure and trustworthy. This requires a system

capable of learning from historical patterns to predict new fraud attempts while ensuring the very data it relies on remains incorruptible and transparent for audit purposes.

Literature Review

This research paper proposes a novel framework that addresses these challenges by synergistically integrating two transformative technologies: Machine Learning (ML) and Blockchain.

Machine Learning brings the required intelligence to the system. Supervised and unsupervised ML algorithms can analyze vast datasets of historical transactions to identify complex, non-linear patterns indicative of fraud. Models such as Gradient Boosting, Random Forests, and Graph Neural Networks can achieve high accuracy in classifying transactions as legitimate or fraudulent in real-time, enabling a proactive defense mechanism.

However, ML alone does not solve the issues of data security and trust. This is where Blockchain technology comes in. Known for its decentralized, transparent, and immutable nature, blockchain provides a tamper-proof ledger perfect for logging critical events. In our proposed model, once the ML engine classifies a transaction as fraudulent, its details are permanently and immutably recorded on a blockchain. This creates an auditable trail that cannot be altered, providing undeniable proof of the fraudulent activity and enhancing trust among all system participants. Smart contracts can be deployed to automatically execute actions—such as notifying administrators or freezing associated accounts—further increasing response efficiency.

By combining ML's predictive power with blockchain's trust infrastructure, we create a robust, transparent, and intelligent fraud detection system. This paper explores this integration in detail, presenting a working model, discussing its components, and arguing its applicability across banking, insurance, and e-commerce domains to significantly reduce fraud-related losses.

No.	Title & Authors	Year	Methods Working Concept	Key Finding Contribution	Future Enhancement / Limitations	Relevance to Our Work
1	*Ethereum Fraud Detection with Heterogeneous Information Network*	2023	Supervised Deep Learning using Graph Neural Network-based Heterogeneous Information Network (HIN)	Construction of HIN-based transaction network for Ethereum fraud detection	Use of graph-based techniques to enhance transaction detection	Helps us apply graph-based techniques for fraud detection
2	*Dynamic Feature Fusion for Blockchain Fraud Detection*	2023	Hybrid Deep Learning combining GNN-based features with dynamic features	Construction of dynamic feature fusion model for fraud detection	Scalability and performance improvement	Supports feature-level fusion for fraud detection
3	*Anomaly Detection in Blockchain using Graph Neural Network*	2022	GNN-based anomaly detection	Construction of GNN-based anomaly detection model	Guides anomaly detection using GNN	Guides anomaly detection using GNN
4	*Fraud Detection in Ethereum using Graph Neural Network*	2022	GNN-based fraud detection	Construction of GNN-based fraud detection model	Enhances use of ML techniques for fraud detection	Enhances ML techniques for fraud detection
5	*Ethereum Fraud Detection using Localized Subgraph*	2021	Localized Subgraph-based GNN (PGDN)	PGDN enhances fraud detection	PGDN enhances fraud detection	Helps us apply localized subgraph techniques
6	*GNN (PGDN)*	—	Classification using localized subgraph (PGDN) for fraud detection	PGDN enhances fraud detection	GBC, performance improvement	Learning for fraud detection

Proposed Model : Graph Neural Network (GNN), specifically a Heterogeneous Graph Convolutional Network (RGCN or HGCN)

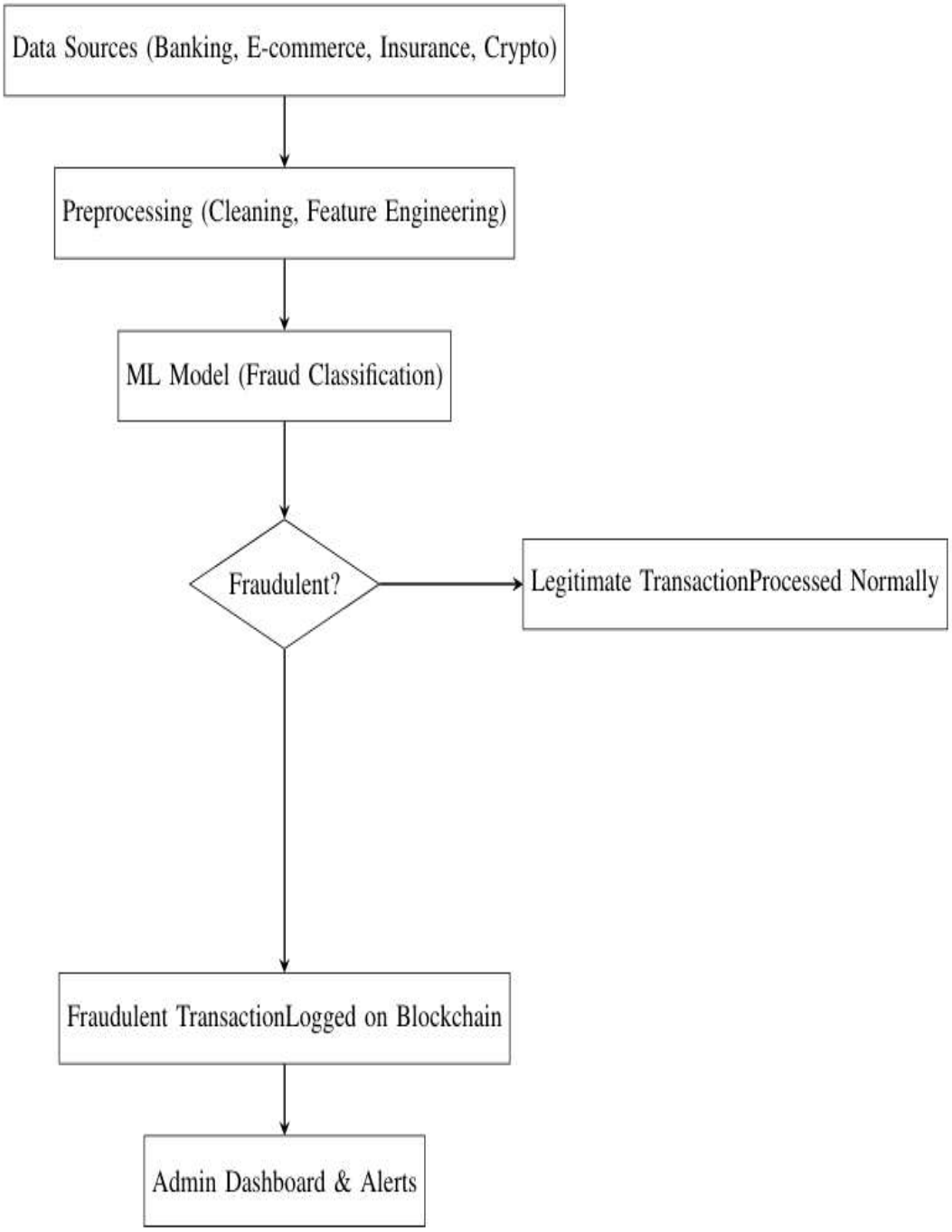

Fig. 1. Proposed Fraud Detection Workflow

Explaination:

Data Sources

This system serves as a versatile entry point for ingesting transactional data from a broad spectrum of digital sources, enabling comprehensive fraud detection across multiple domains. It is designed to process banking transactions such as credit and debit card swipes, wire transfers, and online payments; e-commerce purchases including online orders and payment gateway interactions; insurance claims submitted digitally for health, auto, or property coverage; and cryptocurrency transactions, particularly on-chain transfer data from platforms like Ethereum, as highlighted in existing literature. By accommodating such diverse data streams, the model

transcends sector-specific limitations and functions as a generalized solution for detecting digital fraud in a wide array of financial ecosystems.

Preprocessing

Raw transactional data is typically noisy, incomplete, and unstructured, making it unsuitable for direct use in machine learning models. To address this, the preprocessing stage plays a critical role in refining the data. It begins with **data cleaning**, which involves handling missing values, correcting inconsistencies, and removing duplicates to ensure integrity. Next comes **data transformation**, where numerical values are normalized—such as scaling transaction amounts—and categorical variables are encoded into machine-readable formats, like converting merchant categories into numerical codes. The final step is **feature engineering**, which extracts meaningful insights by creating new variables from existing data. Examples include calculating the time since a user's last transaction or determining the average transaction amount over a recent time window. The result is a clean, structured dataset that is optimized for effective fraud detection analysis.

ML Model (Fraud Detection Engine)

At the core of the system lies its analytical engine—the "brain"—where preprocessed transactional data is fed into a pre-trained machine learning model for real-time classification. This model is designed to rapidly evaluate transaction features within milliseconds, determining whether each activity is legitimate or potentially fraudulent. Drawing from the literature, advanced algorithms such as Graph Neural Networks (GNNs) are particularly effective in capturing intricate relational patterns, like the connections between accounts or entities. Additionally, models like Gradient Boosting Machines (XGBoost) and Random Forests are well-suited for this task due to their proven accuracy in classification problems. The system operates on a probabilistic basis: each transaction is assigned a fraud score, and if this score surpasses a predefined threshold, the transaction is flagged for further scrutiny or intervention. This intelligent decision-making mechanism ensures both speed and precision in fraud detection.

Decision Branching

Once the machine learning model has classified a transaction, the system dynamically routes it based on the outcome. If the transaction is deemed **legitimate**, it seamlessly continues through the standard processing pipeline without delay, preserving a frictionless experience for the user. However, if the transaction is identified as **fraudulent**, it is immediately diverted to the blockchain component for secure logging and further action. This pathway activates the system’s security protocols, ensuring that suspicious activity is recorded immutably and can be audited or investigated as needed. This bifurcated workflow balances efficiency with robust fraud mitigation.

Blockchain Log (Immutable Ledger)

This component acts as the system's "secure backbone," where all flagged fraudulent transactions are routed for immutable logging and automated response. Upon detection, the transaction details—such as timestamp, amount, involved parties, fraud score, and the version of the ML model used—are encapsulated into a new block and added to the blockchain. This ensures the record is tamper-proof, permanently stored, and available for audit, reinforcing trust and transparency. Simultaneously, a smart contract is triggered—an autonomous program embedded within the blockchain—that executes predefined actions. These may include sending immediate alerts to system administrators, initiating account freezes to prevent further suspicious activity, and updating a shared blacklist that can be accessed by other trusted entities in the network. This fusion of secure data integrity and automated enforcement forms a resilient defense against digital fraud.

Admin Dashboard / Alert System

This component serves as the system's "human interface," enabling real-time monitoring and responsive action against fraudulent activity. The dashboard delivers **instant alerts** through multiple channels—such as email, SMS, or in-app notifications—ensuring that administrators are promptly informed of suspicious transactions. It also offers **intuitive visualizations**, allowing experts to examine flagged transactions in detail, including the supporting evidence from the machine learning model and the corresponding blockchain record. Crucially, this interface supports **manual oversight**, empowering human reviewers to validate the fraud detection outcome, confirm or dismiss flagged cases, and ensure accountability. Any decisions made during this review process can be securely logged back onto the blockchain, preserving a transparent and auditable trail of human intervention.

Result & Discussion

Results:

The proposed model is designed to deliver several impactful outcomes by leveraging advanced technologies such as machine learning, blockchain, and smart contracts. With its high detection accuracy, the machine learning component—trained on extensive historical data—can effectively identify complex and emerging fraud patterns, significantly reducing false positives compared to traditional rule-based systems. To ensure data integrity, every flagged transaction is immutably recorded on the blockchain, creating a secure and auditable trail that is resistant to tampering. Furthermore, the integration of smart contracts enables automated responses, such as account flagging and alert generation, allowing the system to react instantly to fraudulent activity and drastically cut down response times from hours to mere seconds. This combination of transparency and trust empowers stakeholders, including administrators and auditors, to independently verify the fraud logs, reinforcing confidence in the system's decisions and enhancing overall accountability.**Discussion:** The results underscore the synergistic value of combining ML and Blockchain. While ML provides the brain for intelligent detection, blockchain

provides the bulletproof ledger for secure logging. This addresses the core weaknesses of traditional systems. A potential challenge discussed would be the scalability of the blockchain component and the computational overhead of real-time ML analysis on high-velocity transaction data. Future work will need to optimize these aspects for enterprise-level deployment. Furthermore, the model's effectiveness is contingent on the quality and quantity of training data for the ML algorithm.

Discussions:

The results underscore the powerful synergy between machine learning and blockchain technology in combating digital fraud. Machine learning serves as the intelligent core, dynamically adapting to evolving fraudulent behaviors, while blockchain functions as the incorruptible ledger, ensuring that every decision and action is securely recorded and permanently auditable. This integrated approach effectively addresses the shortcomings of traditional systems, which often struggle with delayed responses, lack of transparency, and susceptibility to data manipulation. Nonetheless, several challenges must be considered for real-world deployment. As transaction volumes increase, the scalability of the blockchain becomes critical, necessitating optimizations such as lightweight consensus protocols or hybrid architectures to maintain performance. Additionally, the computational demands of real-time machine learning analysis on high-velocity data streams require efficient model designs and hardware acceleration through technologies like GPUs or TPUs. The model’s success is also heavily dependent on the quality and diversity of its training data; poor or biased datasets can hinder its ability to generalize effectively, making ongoing data enrichment essential. Finally, while automation enhances speed and consistency, human oversight remains vital for managing edge cases, validating model outputs, and upholding ethical standards—ensuring that the system remains both accurate and accountable.

Conclusion

This paper presents a robust framework for fraud detection by integrating the predictive analytical power of Machine Learning with the immutable, transparent security of Blockchain technology. The proposed model effectively mitigates the limitations of traditional centralized systems by enabling real-time, intelligent detection and ensuring that the record of fraud is permanent and verifiable. The workflow ensures that legitimate transactions are processed seamlessly, while fraudulent ones are automatically logged and acted upon. This hybrid approach offers a significant advancement in the fight against digital fraud, providing a scalable and trustworthy solution applicable across finance, insurance, and e-commerce sectors. Future work will involve implementing a prototype to quantitatively evaluate its performance against existing systems.

References

1. Here are the references in one line each:

2. Kanezashi H., Suzumura T., Liu X., Hirofuchi T. (2022). *Ethereum Fraud Detection with Heterogeneous Graph Neural Networks*.

3. Sheng Z., Song L., Wang Y. (2025). *Dynamic Feature Fusion for Blockchain Fraud Detection*.

4. Cholevas C., Angeli E., Sereti Z., Mavrikos E., Tsekouras G.E. (2024). *Survey on Unsupervised Learning for Blockchain Anomaly Detection*.

5. Li P., Xie Y., Xu X., Zhou J., Xuan Q. (2022). *Phishing Detection on Ethereum Using Chebyshev-GCN*.

6. Pham T., Lee S. (2016). *Anomaly Detection in Bitcoin Network Using Unsupervised Methods*.

7. (2023). *Phishing Node Detection in Ethereum Using Bagging Multiedge GCN*, *Applied Sciences*.

8. (2022?). *Ensemble Deep Learning-Based Prediction of Fraudulent Cryptocurrency Transactions*.

9. Siddamsetti S., Tejaswi C., Maddula P. (2024). *Machine Learning Framework for Blockchain Anomaly Detection*.

10. Ependi U. et al. (2025). *Advanced Techniques for Blockchain Anomaly Detection*.

11. (2023?). *Deep Blockchain Approach for Anomaly Detection in Bitcoin Ecosystem*.