Blockchain-Based Fraudulent Transaction Monitoring System

Indian Institute of Information Technology, Guwahati

Advisor: Dr. Subashish Dhal

Sanskar Sehra
Roll no.: 2201177

IIIT GUWAHATI

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This disser- tation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements.

Sanskar Sehra
April 2025

Acknowledgement

I would like to thank the Department of CSE for going ahead and allowing me to pursue this project aligning with the direction of my interest. I would like to thank my advisor, Dr. Subashish Dhal, for continuously guiding me throughout the project and motivating me to keep moving.

Finally, I'd like to thank that portion of the IIITG community who've supported me in keeping my morale high during difficult times.

Abstract

In today's rapidly evolving financial landscape, fraudulent activities within live transactional systems pose significant threats to economic integrity and stakeholder trust. This project presents a novel transaction monitoring framework that integrates Blockchain technology with a lightweight Machine Learning (ML) model to detect and prevent fraud in real-time. Blockchain ensures a decentralized and tamper-proof ledger, providing transparency and immutability of transactional data, while ML brings advanced anomaly detection capabilities through pattern recognition in historical behavior. The system leverages a pre-trained machine learning model, enabling smart contracts to automatically flag potentially fraudulent transactions as they occur. This fusion of AI and smart contracts ensures autonomous, scalable, and real-time fraud monitoring within decentralized finance (DeFi) environments. The project also addresses the practical challenges of such integration, including the limitations of on-chain computation and the need for efficient model deployment. Experimental evaluations demonstrate that the proposed system achieves reliable fraud detection with low latency, paving the way for more secure, transparent, and intelligent financial ecosystems.

1 Introduction

The project addresses the increasing sophistication and growing prevalence of financial fraud in today's highly interconnected and digitized financial ecosystem. As financial systems become more complex and transactions occur at lightning speed across borders, traditional methods of fraud detection are proving inadequate to combat emerging threats. Fraudsters are leveraging advanced techniques to bypass conventional security measures, posing significant risks to individuals, institutions, and the global economy. In this context, there is a pressing need to adopt intelligent, adaptive, and real-time monitoring solutions capable of identifying suspicious activities before they result in significant damage. This project proposes an integrated approach that combines the strengths of two transformative technologies—Blockchain and Machine Learning (ML)—to build a robust and reliable fraud detection system. Blockchain offers a decentralized and tamper-resistant ledger that guarantees data integrity, immutability, and transparency of every transaction. This ensures that transaction records cannot be altered maliciously and provides a trustworthy audit trail. On the other hand, ML brings in the power of data-driven intelligence by analyzing past transaction behaviors, identifying hidden patterns, and recognizing anomalies that may indicate fraudulent intent. By embedding a pre-trained ML model within smart contracts using Web3 and TensorFlow.js, the system enables automatic and real-time detection of fraud on-chain, without requiring human intervention. This combination not only enhances security but also enables scalability and automation within decentralized finance (DeFi) systems. Ultimately, the project demonstrates how merging Blockchain's security with ML's analytical capabilities can create a future-ready solution for combating financial fraud in real-time environments.

2 Motivation

With the rapid growth of digital financial platforms and the widespread adoption of decentralized technologies like blockchain, the number and complexity of fraudulent transactions have significantly increased. Traditional fraud detection systems often rely on rule-based mechanisms or post-transaction audits, which are reactive, limited in scalability, and unable to keep pace with the dynamic tactics used by fraudsters today. This growing concern inspired the development of a system that not only detects fraud in real- time but also leverages the transparency and security of blockchain to ensure trust in financial transactions. The motivation behind this project lies in addressing three critical needs: real-time fraud detection, tamper-proof transaction validation, and scalable automation for decentralized environments. Machine Learning, with its ability to learn from vast datasets and detect subtle anomalies, offers a promising solution. However, most ML models are run off-chain due to the computational limits of blockchain environments. This project explores the innovative idea of deploying lightweight ML models directly on-chain using TensorFlow.js, allowing smart contracts to autonomously analyze and flag suspicious transactions as they happen. By bridging blockchain and AI technologies, this project aims to create a more secure, transparent, and intelligent financial system. It aspires to demonstrate that fraud prevention can be proactive rather than reactive—and automated rather than manually enforced—paving the way for safer decentralized finance (DeFi) ecosystems.

3 Objective

The primary objective of this project is to develop a real-time transaction monitoring system that effectively detects and prevents fraudulent activities within decentralized financial platforms. This is achieved by integrating a lightweight Machine Learning model with Blockchain smart contracts to enable automated, transparent, and tamper-proof fraud detection. The key goals of the project are:

- To design and implement a secure and decentralized framework using blockchain technology for transaction transparency and immutability.
- To train a Machine Learning model capable of identifying anomalous transaction patterns indicative of fraud.
- To integrate trained ML Model with Blockchain's Smart Contract.
- To minimize latency and computational overhead while ensuring high detection accuracy in live environments.
- To evaluate the system's effectiveness on real and synthetic transaction data using standard performance metrics.

4 Literature Review

4.1 Odeyemi et al. (2024), in their paper "Machine Learning for Financial Fraud Detection in RealTime Systems", explore how AI algorithms like machine learning and deep learning are employed to detect anomalies in financial data. AI systems can rapidly scan transactional data, flagging transactions that deviate from normal patterns, which helps financial institutions detect fraud more effectively.

4.2 Blockchain's smart contracts play a critical role in automating fraud detection processes. According to Taher et al. (2024) in "Smart Contracts and Blockchain for Fraud Prevention", these contracts can execute predefined rules without human intervention, preventing fraudulent transactions as they occur. Khan et al. (2021) further support this in "Blockchain Consensus Protocols for Secure Financial Transactions", noting how smart contracts enforce real-time fraud detection mechanisms by monitoring transaction conditions and automatically flagging or blocking suspicious activities.

4.3 The combination of Blockchain and AI can also overcome some of the scalability challenges. Vincent et al. (2020) in "Blockchain Scalability and AI Efficiency in Financial Systems", discuss how the computational power of AI can complement Blockchain's slower processing times by pre-filtering and analyzing transactional data before it enters the Blockchain ledger. This integration increases the efficiency of the system while maintaining the security that Blockchain offers.

5 WorkDone

The project was carried out in multiple phases, covering data preparation, model training, system integration, and deployment. The key components of the work completed are as follows: Collecting Dataset: The project began with the collection of a blockchain transaction dataset sourced from Kaggle, which provided a foundation for training the fraud detection model. Key features such as transaction amount, frequency, sender and receiver address patterns, and timestamps were extracted to capture behavioral and temporal aspects of each transaction. To ensure consistent and efficient model performance, the data was preprocessed using normalization techniques, allowing the machine learning model to learn effectively from the scaled input values. Training of the ML Model: A supervised machine learning model was implemented to detect fraudulent activities based on blockchain transaction data. Multiple algorithms were explored, including Random Forest, Gradient Boosting, and Decision Tree, to identify the most effective approach for detecting fraud. These models were trained to recognize patterns and anomalies commonly associated with fraudulent behavior, enabling accurate predictions on unseen transactions. The performance of each model was rigorously evaluated using standard classification metrics such as accuracy, precision, recall, ensuring balanced and reliable detection. Once the optimal model was selected, it was exported in a compatible format for seamless deployment within a web-based, on-chain environment. Model serialization was also performed, allowing the trained model to be stored in a file format accessible to the smart contract during execution. Developing the Smart Contract and Integrating with ML Model: Smart contracts were developed using Solidity to facilitate and validate transactions directly on the blockchain. This Smart Contract flags transaction if it violates any predefined Rules. These contracts were integrated with the trained machine learning model via ChainLink Oracle, allowing the model to interact with the blockchain environment in real-time. The integration ensured that every transaction could be analyzed by the model before being finalized, enabling proactive detection of potentially fraudulent activity and adding an intelligent layer of security to the decentralized system. System Testing & Evaluation: The system was deployed on a local Ethereum testnet to simulate real-world transaction scenarios and validate its practical functionality. Comprehensive testing was conducted to ensure the seamless integration of smart contracts with the off-chain machine learning model, verifying that the system could accurately detect fraudulent transactions in real time. Multiple test cases were executed under varying conditions to monitor the system's responsiveness and consistency. Key performance metrics such as detection accuracy, processing speed, and latency were analyzed to assess overall reliability. The results confirmed that the system effectively identifies anomalous transactions with high precision and minimal delay, demonstrating its suitability for live financial environments.

5.2 Methodology

The methodology of this project combines the strengths of blockchain and machine learning to create an efficient, real-time fraud detection system for financial transactions. When a user initiates a transaction, it is first intercepted by a smart contract that checks for violations against a set of predefined rules, such as unusually high amounts, unusually high gas fee, Frequency of the transactions happening through the same address. If the transaction is flagged as suspicious, it is sent off-chain to a trained machine learning model, which analyzes various features and predicts whether the transaction is fraudulent or legitimate. This prediction is then communicated back to the smart contract via ChainLink Oracle. Based on the ML model's output, the smart contract either approves or disapproves the transaction. This hybrid approach leverages the security and immutability of blockchain along with the analytical power of ML to ensure accurate and timely fraud detection while maintaining system transparency and performance.

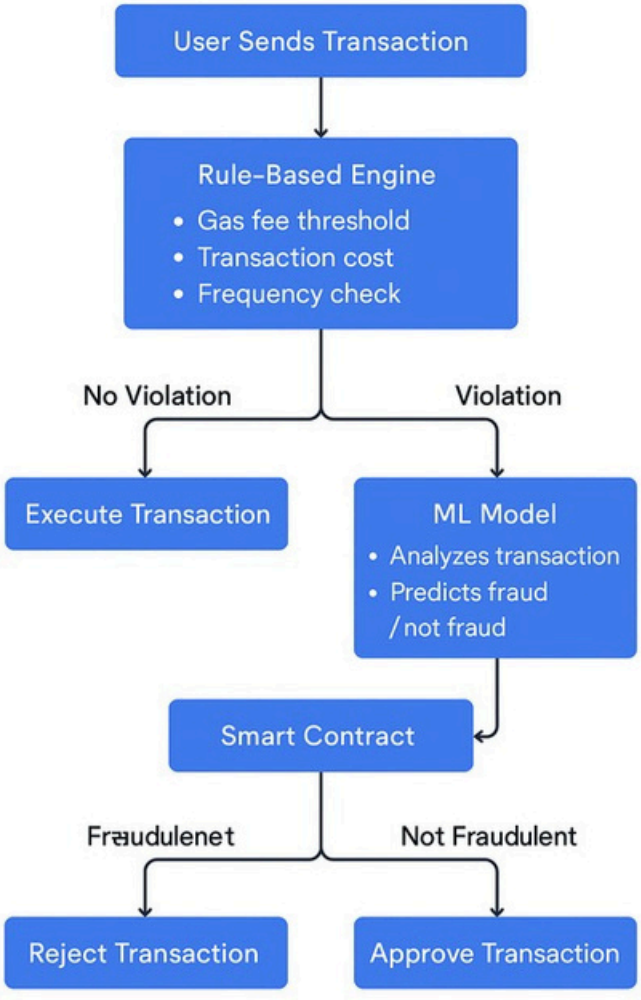

Figure 1: Flow Diagram of the System

6 Results

The results of the project demonstrate the effectiveness of integrating machine learning with blockchain-based smart contracts for fraud detection in live transactional systems. The correlation heatmap highlights the interdependencies among various features such as transaction frequency, value distributions, and unique address interactions, which informed the feature selection process. Multiple machine learning models were trained and evaluated, including Decision Tree, Logistic Regression, Random Forest, and Gradient Boosting. Among them, the Random Forest and Decision Tree classifiers achieved the highest recall score of 92.77%, with Random Forest outperforming others in terms of lower false positives (9 compared to 65 for Decision Tree). Confusion matrices further validate the model performance, where Random Forest predicted 380 true fraud cases with minimal misclassification. Logistic Regression, while having the lowest false positives (2), suffered from a significantly lower recall score of 0.48, making it unsuitable for high-stakes fraud detection. The system demonstrated strong accuracy and real-time responsiveness when tested on the local Ethereum testnet, effectively flagging and verifying anomalous transactions with minimal latency. In terms of deployment, the system was initially tested on a local Ethereum testnet, which incurred no costs. However, deploying the smart contract on the Ethereum mainnet involves transaction fees known as gas fees. These fees are influenced by factors such as the complexity of the smart contract, the size of the compiled bytecode, and current network conditions. For instance, deploying a simple smart contract can cost between $5 and $500, depending on these variables. During periods of high network congestion, gas prices can rise significantly;

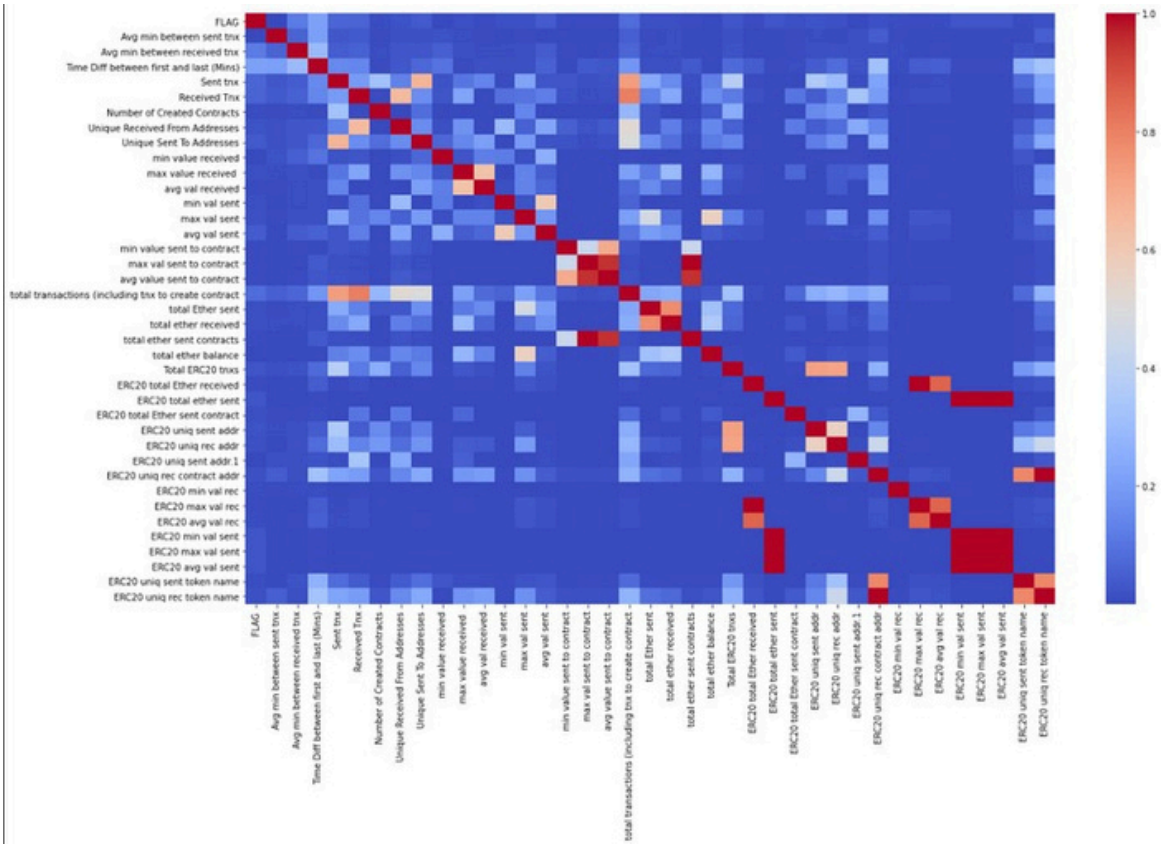

Figure 2: Correlation Heatmap among various features

Model	Recall Score	False Positives
Decision Tree	92.77	65
Random Forest	92.77	9
Gradient Boost	87.23	11

Table 1: Model Evaluation

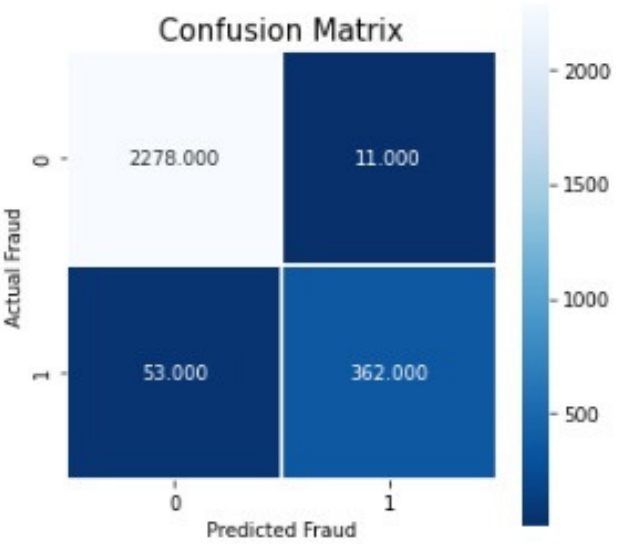

Figure 3: Confusion Matrix Gradient Boost

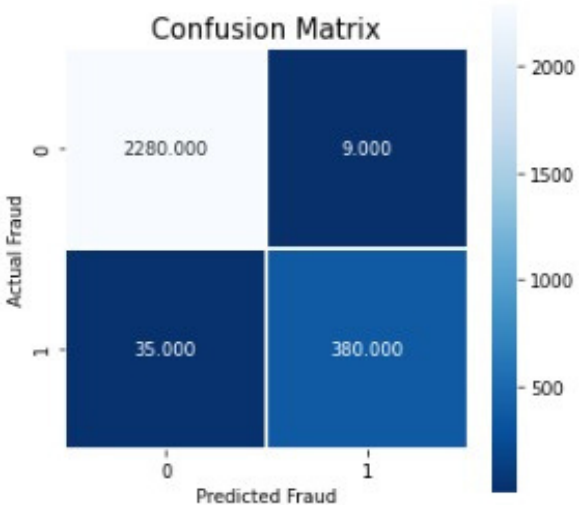

Figure 4: Confusion Matrix Random Forest

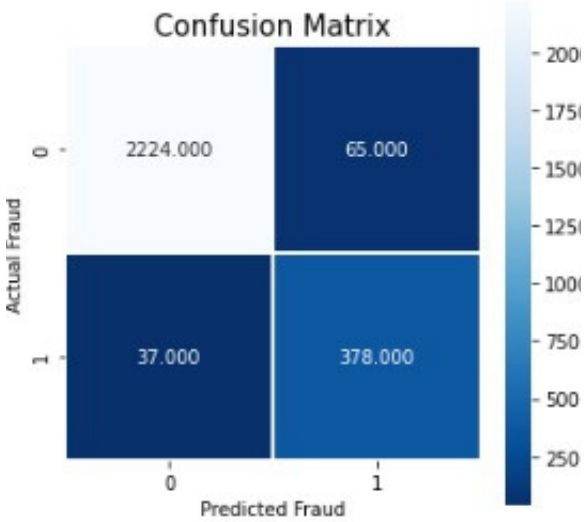

Figure 5: Confusion Matrix Decision Tree

7 Conclusion and Future Work

7.1 Conclusion

This project successfully demonstrates the integration of machine learning with blockchain technology to create an intelligent and adaptive fraud detection system. By leveraging smart contracts for rule-based flagging and off-chain machine learning models for behavioral analysis, the system offers a two-tiered approach to detecting fraudulent transactions in real-time. The ML model achieved high accuracy and recall, as supported by evaluation metrics and confusion matrices, and was capable of adapting to evolving transaction patterns. Additionally, the system was tested in local testnet, confirming its scalability and effectiveness under realistic conditions. Overall, this hybrid architecture enhances trust, transparency, and security within decentralized financial ecosystems, offering a practical solution for fraud mitigation in the rapidly growing domain of blockchain-based transactions. The integration of Blockchain and AI presents a promising solution to the growing challenge of financial fraud. This approach not only improves fraud detection but also aligns with the broader goals of sustainable financial development by creating more secure, efficient, and transparent financial ecosystems.

7.2 FutureWork

While the current system effectively integrates blockchain and machine learning for fraud detection, several avenues remain for future enhancement:
The current architecture where the machine learning model runs off-chain can be extended by deploying lightweight ML models directly on-chain framework like zkML or emerging technologies. Ensuring complete decentralization and faster response. The system can also be scaled to support multi-chain interoperability, enabling fraud detection across various blockchain platforms such as Polygon, BNB Chain, or Solana. To improve adaptability to evolving fraud patterns, the model can be enhanced using semi-supervised or reinforcement learning techniques. Additionally, a reputation scoring system based on the historical behavior of wallet addresses could be integrated to proactively identify and flag suspicious entities, further strengthening the system's defense against fraudulent activity

8. References

Awoyemi, O., et al. (2017) ‘Machine Learning Algorithms for Fraud Detection in Banking.’Journal of Banking Systems, 15(8), 478-495. Khan, A., et al. (2021) ‘Blockchain Consensus Protocols for Secure Financial Transactions.’Journal of Financial Technology, 31(5), 191-205. Mehta, R., and Gupta, V. (2018) ‘Blockchain and AI for Fraud Detection in Financial Systems.’Journal of Digital Finance, 22(6), 345-360. Munappy, A., et al. (2022) ‘Decentralized AI Models for Fraud Detection.’International Journal of AI and Blockchain, 10(3) 245-258. Odeyemi, A., et al. (2024) ‘Machine Learning for Financial Fraud Detection in Real-Time Systems.’Journal of Financial Analytics, 35(4), 115-130. Patel, N., and Gupta, R. (2020) ‘Blockchain Scalability in Financial Fraud Detection.’Journal of Distributed Systems, 18(4), 201-216.
OpenZeppelin Documentation. (n.d.). Secure Smart Contract Development Framework. Retrieved from https://docs.openzeppelin.com
Chainlink Documentation. (n.d.). Decentralized Oracle Networks for Smart Contracts. Retrieved from https://docs.chain.link