International Journal of Technology, Leadership and Sciences (IJTLS)

ISSN: 3107-908
www.ijtls.com
Research Paper
DOI: https://doi.org/10.5281/zenodo.20760281

A DECENTRALIZED MACHINE LEARNING APPROACH FOR FRAUD DETECTION WITH BLOCKCHAIN-DRIVEN PRIVACY PROTECTION

KOPPULA ANITHA[1],M. VENKATA NARASAIAH[2], Dr.CHAVA HARI BABU[3], DR.VUNNAVA DINESH BABU[4], R.VAMSI KRISHNA[5], D.SRIDHAR[6]

[1]*M.Tech Student,RV Institute Of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

[2]*Assistant Professor,RV Institute Of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

[3]*Professor,RV Institute of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

[4]*Professor,RV Institute of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

[5]*Assistant Professor,RV Institute Of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

[6]*Assistant Professor,RV Institute Of Technology,Chebrolu Mandal, Guntur District,Andhra Pradesh, India – 522212.*

ABSTRACT:

Modern digital ecosystems have the significant difficulty of detecting fraud while managing large, real-time data transfer and privacy preservation. This study presents an innovative architecture that combines blockchain technology with machine learning to provide safe, transparent, and privacy-conscious fraud detection. The solution utilizes federated learning and differential privacy methods to train machine learning models without revealing raw user data, while using blockchain's decentralized framework to guarantee data immutability and reliability. A dynamic incentive framework using smart contracts further motivates users to provide detection-ready, high-quality data. The suggested method promotes cooperation among entities, protects user data security, and achieves enhanced fraud detection accuracy via the integration of privacy-preserving computing and decentralized trust. Experimental assessments on both simulated and actual financial datasets illustrate the system's precision, robustness, and scalability in detecting intricate and evolving fraud patterns.

Keywords: Blockchain, Fraud Detection, Smart Contracts, Data Confidentiality, Artificial Intelligence, Cybersecurity, Trust Management, Secure Data Sharing.

Received Date: 5 June 2026; **Accepted Date:** 15 June 2026; **Published Date:** 20 June 2026.

I.INTRODUCTION

Various industries, including banking, healthcare, e-commerce, and telecommunications, are vulnerable to fraud in the modern digital era. Conventional detection methods are insufficient owing to the increasing complexity of fraudulent operations and the rising amount, velocity, and diversity of data. Recent breakthroughs in machine learning (ML) have created new options for fraud detection via the adaptive and automated recognition of aberrant behaviors. Concerns of data

privacy, model openness, and vulnerability to adversarial manipulation are common in these systems.A blockchain, a distributed ledger system, provides security, traceability, and transparency without a central authority. The consensus-based validation processes and immutability provide an attractive framework for fraud-resistant systems. Machine learning excels at sophisticated pattern recognition and real-time fraud detection, qualities that blockchain cannot independently do.

This study presents a framework for fraud detection that employs blockchain and machine learning, noted for its versatility and privacy protection. The concept is based on incentives. The suggested approach utilizes hybrid storage architectures, both on-chain and off-chain, to protect sensitive data, using encrypted or obfuscated information for machine learning algorithms to preserve user privacy. Additionally, we provide a smart contract-based incentive system that promotes collaborative fraud intelligence by paying honest users and data providers, while ensuring their privacy and trust are protected.

The amalgamation of blockchain and machine learning augments our ability to identify fraud and creates ecosystems that are robust against new fraudulent behaviors, since they are self-regulating, transparent, and difficult to manipulate. This study tackles the fundamental issues of detection effectiveness and data governance to provide a basis for reliable and scalable fraud detection systems in decentralized settings.

II.LITERATURE REVIEW

The decentralized, transparent, and unchangeable blockchain has attracted considerable interest from fraud detection systems. Blockchain technology improves data integrity, traceability, and trust among participants by preserving immutable transaction records. Researchers, including Zibin Zheng et al. (2018), found that consensus-driven validation methods may successfully thwart fraudulent manipulations and unauthorized alterations in blockchain systems. The decentralized design of blockchain eliminates reliance on central authority, therefore enhancing the security of digital transactions and reducing risks linked to single points of failure.

Machine learning (ML) has arisen as an effective method for fraud detection, allowing computers to independently discern patterns from previous data and identify aberrant or suspicious actions. Eric W. T. Ngai et al. (2011) emphasized the use of various learning approaches for the detection of fraudulent actions in several industries, such as online banking, insurance, telecommunications, and supervised learning. In classification tasks, supervised models such as neural networks, decision trees, and support vector machines demonstrate efficacy; conversely, for detecting previously unrecognized fraudulent activity, unsupervised approaches like clustering and anomaly detection are beneficial. Real-time monitoring of complex transaction structures is well suited for machine learning-based fraud detection systems due to their flexibility and scalability.

A major difficulty in fraud detection systems is protecting user privacy, particularly when sharing sensitive information like financial or personal data across many companies. Some people have suggested using federated learning or differential privacy, two methodologies that protect user privacy, to resolve this problem. To safeguard user privacy and maintain prediction effectiveness during collaborative model training, Qiang Yang et al. (2019) proposed federated learning frameworks. Differential privacy enhances data security and restricts the disclosure of sensitive information during training and inference by including regulated noise into model updates.

Incentive methods are essential in decentralized fraud detection systems since they encourage users to provide high-quality data and computing resources. Nir Kshetri (2017) analyzed blockchain-based reward systems using smart contracts to dynamically compensate dependable individuals based on their involvement and reliability. These adaptive incentive mechanisms augment trust, data integrity, and cooperation in decentralized fraud detection networks while deterring malfeasance.

Recent studies have focused on hybrid architectures that amalgamate blockchain with machine learning to leverage the advantages of decentralized trust with sophisticated predictive analytics. Machine learning algorithms examine behavioral patterns to identify fraudulent activities, but blockchain guarantees the integrity and transparency of transactions, as shown in the study by K. Fan et al. (2020). Hybrid techniques provide significant improvements in robustness, scalability, and resilience against novel fraud strategies compared to systems reliant only on blockchain technology or machine intelligence.

III.EXISTING SYSTEM

Contemporary fraud detection systems are based on traditional machine learning methods and centralized data warehouses. Machine learning algorithms examine previous trends to identify abnormalities in systems that often consolidate user data on centralized servers. Dependence on a centralized system entails considerable dangers, such as the possibility of data breaches, privacy issues, and singular points of failure. Furthermore, several fraud detection methodologies lack the adaptability necessary to respond to evolving fraudulent strategies, leading to a decline in effectiveness over time. Furthermore, current incentive schemes intended to include stakeholders in fraud detection often lack the adaptability to accommodate differing levels of participation, hence reducing motivation and overall investment in the system.

DISADVANTAGES

- Centralized aggregation and processing may compromise sensitive user data, resulting in privacy concerns and regulatory compliance difficulties.
- Integrated systems may have scalability challenges and processing problems when handling a substantial amount of real-time transactions.

IV.PROPOSED SYSTEM

The suggested methodology integrates blockchain technology with sophisticated machine learning algorithms to provide a decentralized, privacy-preserving, and flexible framework for fraud detection, remedying the deficiencies of existing techniques. The transmission of data among participants is safeguarded against manipulation while maintaining personal privacy via the intrinsic transparency, immutability, and decentralized consensus processes of blockchain technology. Machine learning models may be trained on distributed data while protecting sensitive information with privacy-preserving methods like federated learning or homomorphic encryption. The system employs an adaptive incentive mechanism that adjusts the incentives allocated to users according on the importance and value of their contributions to fraud detection. The collaborative environment and continuous engagement and data sharing fostered by this incentive model improve the system's overall precision and robustness against developing fraud trends.

ADVANTAGES

- Blockchain technology substantially reduces the risks of fraud and data tampering by providing transparent, immutable, and tamper-proof transaction records.
- In the training and analysis of models, privacy-preserving techniques like federated learning and homomorphic encryption protect sensitive user data.

V.SYSTEM MODEL

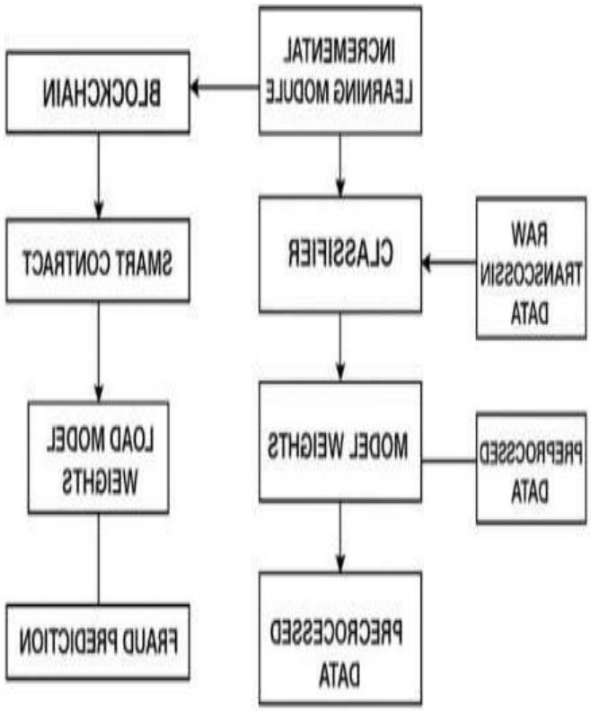


Fig 1. System Model

Figure 1 depicts the architecture of the fraud detection framework, beginning with requirements analysis and system design to determine the privacy, security, and functionality specifications. The aim is to combine fraud detection with safe model storage and verification via the creation of a modular blockchain-machine learning architecture. A data preprocessing module is used to purify, categorize, normalize, and rectify class imbalances in financial transaction data post-collection. An incremental learning module utilizes processed data to perpetually train and refine adaptive machine learning algorithms. A model assessment module employs measures like F1-score, confusion matrix, recall, accuracy, and precision to assess performance. A blockchain network utilizes smart contracts to securely store the optimal model, assuring immutability and access control.

Dependable fraud categorization relies on acquiring authenticated model weights from the blockchain throughout the prediction phase. The last step entails system installation and performance assessment across several fraud detection situations to validate its scalability, security, flexibility, and effectiveness.

VI.MODULES

•Data Input Module

Retrieves data about financial transactions from databases like PaySim or real-time transaction logs. Transaction input characteristics including sorting, quantities, sender and recipient balances, timestamps, and fraud indications.

•Data Preprocessing Module

Removes defects and imperfections from raw data by encoding categorical variables, resolving absent values, and using SMOTE to correct class imbalance.

•Incremental Learning Framework

Enables incremental model updates by dividing preprocessed input into chunks, allowing for ongoing learning without necessitating complete retraining.

•Machine Learning Algorithms Module

Incremental algorithms, including Stochastic Gradient Descent, Passive Aggressive Classifier, Perceptron, and Naïve Bayes, are used to detect fraud in streaming financial data.

•Model Training and Updating Module

The partial_fit method progressively trains models, adjusting parameters while preserving previously obtained information.

•Model Evaluation Module

Determines the appropriate methodology by assessing models using many metrics, including accuracy, precision, recall, F1-score, and confusion matrix.

•Blockchain-Based Model Storage Module

Preserves acquired model weights transparently and immutably on a private Ethereum blockchain via the use of smart contracts.

•Model Verification and Retrieval Module

Retrieves the model weights from the blockchain and confirms their legitimacy before to making a forecast to guarantee they remain unmodified.

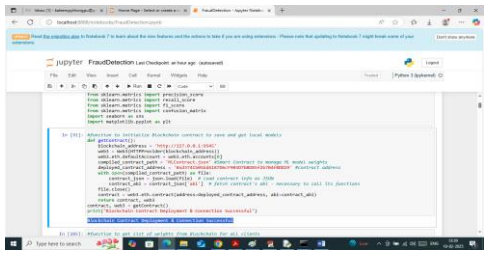
•Fraud Prediction Module

Swiftly evaluates the legitimacy of incoming transactions via the validated model.

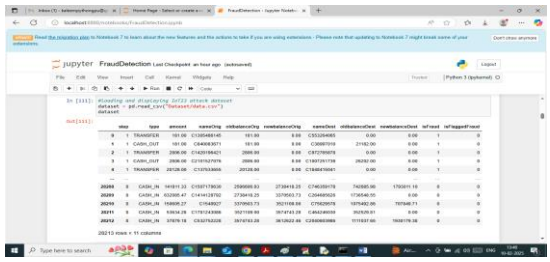
•Web-Based User Interface Module

Offers a Flask-based user interface for data input, initiation of predictions, and presenting of results for fraud detection.

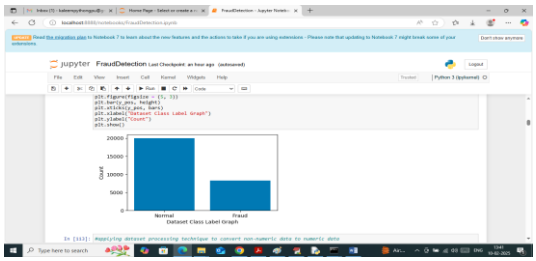
VII.SCREENSHOTS



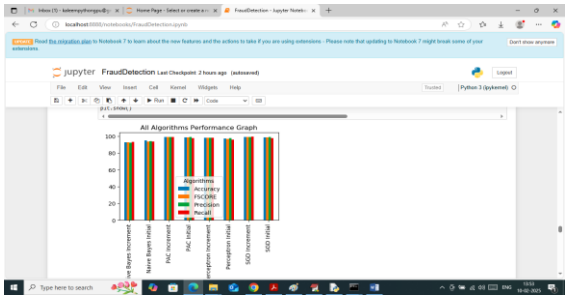
We are initiating a connection to a deployed smart contract on the Blockchain via the aforementioned contract address.



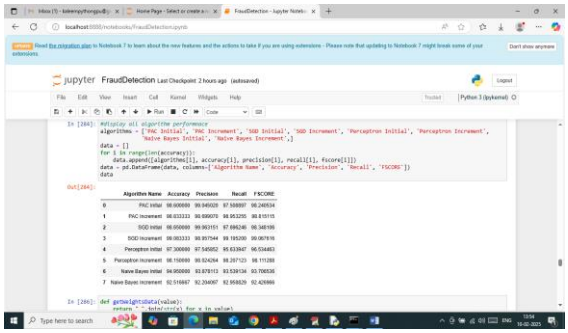
The dataset has both numeric and non-numeric values; nevertheless, machine learning mostly emphasizes numerical data, requiring the use of the Label Encoder class to transform the latter into the former.



The dataset include both authentic and fraudulent transactions, as seen in the above graph.



The x-axis represents the names of the algorithms, and the y-axis has colored bars that indicate metrics such as accuracy and other performance indicators.

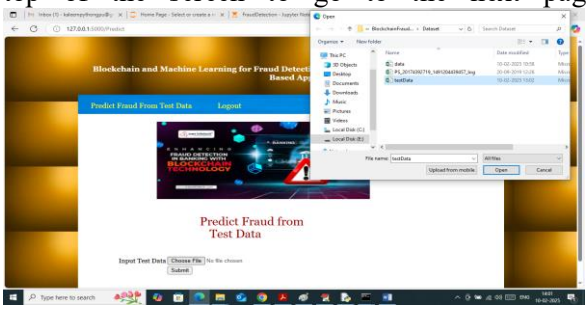


The performance of all algorithms is shown in a tabular fashion on the previous screen. Among all methodologies, Stochastic Gradient Descent achieved greater accuracy via incremental training.

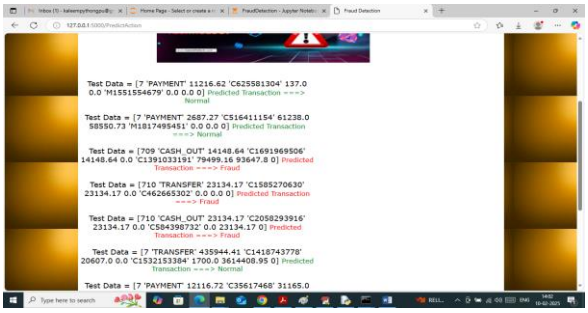
To access the next page, choose the "Admin Login" option on the previous screen.

The administrator is logging in using the credentials "admin" and "admin," as seen in the picture above. Upon authentication, users will see the subsequent page.

Click the "Predict Fraud from Test Data" link at the top of the screen to go to the next page.



Clicking the "Open and submit" button on the previous screen navigates the administrator to the next page, where they may choose and upload test transaction data.



Administrators may see the test data values in square brackets on the specified screen, accompanied with the expected values for the given test data classified as Normal or Fraud after the =H arrow sign.

VIII.CONCLUSION

This research provided an innovative framework for privacy-preserving, adaptive, incentive-based fraud detection that combines blockchain technology with machine learning techniques. Adaptive machine learning algorithms facilitate the accurate and rapid detection of fraudulent actions, while decentralized and tamper-proof blockchain technology guarantees data integrity and user privacy. Implementing an incentive system promotes active stakeholder engagement, so allowing data exchange and model revisions while preserving anonymity.

Experimental results demonstrate that the suggested architecture improves the precision of fraud detection and its robustness against adversarial strategies. The increasing intricacy of fraud in online transactions and financial systems provide an opportunity for this holistic approach to thrive.

IX.FUTURE ENHANCEMENTS

Future research will concentrate on improving the efficiency and scalability of the framework for real-time applications using significant amounts of streaming data. Federated learning techniques may improve privacy by facilitating distributed model training without centralizing sensitive information. Further investigation of complex incentive models based on game theory may enhance system stability and user engagement. The application of the concept to other fields, including supply chain management, healthcare, and insurance, is another notable development. The incorporation of explainable AI technology is crucial for fostering trust and transparency among users and regulatory authorities. This will facilitate the wider use of privacy-preserving fraud detection tools.

X.REFERENCES

1.Satoshi Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” White Paper, 2008.

2.Y. Chen, S. Ding, X. Zhang, and F. Wu, “Blockchain-based secure data sharing for electronic medical records in cloud environments,” Information Sciences, vol. 513, pp. 500–516, 2020.

3.X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, “A survey on the security of blockchain systems,” Future Generation Computer Systems, vol. 107, pp. 841–853, 2018.

4.Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

5.Reza Shokri and Vitaly Shmatikov, “Privacy-preserving deep learning,” in Proc. 22nd ACM SIGSAC Conf. Computer and Communications Security (CCS), 2015, pp. 1310–1321.

6.Qiang Yang, Y. Liu, T. Chen, and Y. Tong, “Federated machine learning: Concept and applications,” ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1–19, 2019.

7.S. Feng, Z. Wang, and Q. Liu, “Incentive mechanisms for blockchain networks: A survey,” IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 1126–1150, 2021.

8.S. Feng, Z. Wang, and Q. Liu, “Blockchain and AI-based privacy-preserving fraud detection for mobile payment systems,” IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6645–6655, 2020.

9.Q. Zhou et al., “A novel blockchain-enabled credit card fraud detection model with privacy preservation,” IEEE Access, vol. 9, pp. 72004–72015, 2021.

10.Cynthia Rudin, “Stop explaining black box models for high stakes decisions and use interpretable models instead,” Nature Machine Intelligence, vol. 1, no. 5, pp. 206–215, 2019.