ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

BLOCKCHAIN AND MACHINE LEARNING FOR FRAUD DETECTION: A PRIVACY-PRESERVING AND ADAPTIVE INCENTIVE BASED APPROACH

MRS. TALARI KALPANA

M. TECH CSE DEPARTMENT, SANSKRITHI SCHOOL OF ENGINEERING, PUTTAPARTHY,SRISATYA SAI DIST, ANDHRAPRADESH,, INDIA,

KALPANAPOTTI1804@GMAIL.COM@GMAIL.COM

GUIDE

MR K. VINOD KUMAR

CSE DEPARTMENT, SANSKRITHI SCHOOL OF ENGINEERING, PUTTAPARTHY,SRISATYA SAI DIST, ANDHRAPRADESH, INDIA

ABSTRACT

Fraud detection has become increasingly important in the digital era due to the rapid growth of online transactions, financial technologies, and e-commerce platforms. Traditional fraud detection systems are primarily centralized and rule-based, which makes them vulnerable to data breaches, limited scalability, and inability to detect evolving fraud patterns. These systems also face challenges related to data privacy, as sensitive financial and personal information is often stored and processed in centralized databases. To address these limitations, this paper proposes a novel framework that integrates Blockchain technology with Machine Learning to develop a secure, decentralized, and privacy-preserving fraud detection system.

The proposed system leverages blockchain to provide a distributed and immutable ledger, ensuring transparency, trust, and data integrity among participants. Machine learning algorithms are used to analyze transaction data and identify suspicious activities by detecting anomalies and patterns indicative of fraud. Instead of sharing raw data, the system utilizes privacy-preserving techniques such as distributed or federated learning, where only model parameters or encrypted insights are shared across the network.

Additionally, the framework introduces an adaptive incentive-based mechanism to encourage participants to contribute high-quality data and computational resources. This mechanism ensures active participation and improves model performance over time. Experimental evaluations demonstrate that the integration of blockchain and machine learning significantly enhances fraud detection accuracy while maintaining data privacy and system reliability. The proposed approach is scalable, secure, and suitable for real-world applications in banking, insurance, and digital payment systems.

Keywords: Blockchain, Machine Learning, Fraud Detection, Privacy Preservation, Incentive Mechanism, Cybersecurity, Distributed Systems.

1. INTRODUCTION

The increasing digitization of financial services and online platforms has significantly transformed the way transactions are conducted. While this transformation has improved efficiency and accessibility, it has also led to a rise in fraudulent activities such as credit card fraud, identity theft, insurance fraud, and cyberattacks. These fraudulent activities result in substantial financial losses and pose serious threats to individuals, organizations, and governments. Traditional fraud detection systems rely on centralized databases and predefined rules, which are often insufficient to detect sophisticated and evolving fraud patterns.

Machine Learning (ML) has emerged as a powerful tool for fraud detection due to its ability to analyze large volumes of data and identify hidden patterns. ML models can learn from historical transaction data and detect anomalies that may indicate fraudulent behavior. Techniques such as decision trees, neural networks, and ensemble methods have shown promising results in improving detection accuracy. However, these models require access to large datasets, which raises concerns about data privacy and security.

Blockchain technology offers a decentralized and secure approach to data management. It provides a distributed ledger where transactions are recorded in a transparent and immutable manner. By integrating blockchain with machine learning, it is possible to create a fraud detection system that ensures data integrity and privacy. Furthermore, the inclusion of an adaptive incentive mechanism motivates participants to contribute data and computational resources, enhancing system performance.

This paper proposes a hybrid framework that combines blockchain and machine learning to address the challenges of fraud detection. The system aims to provide a secure, scalable, and privacy-preserving solution that adapts to evolving fraud patterns while maintaining trust among participants.

2. LITERATURE SURVEY

Fraud detection has been widely studied using various computational techniques, particularly machine learning and data mining approaches. Early research focused on rule-based systems that relied on predefined conditions to identify fraudulent transactions. However, these systems lacked adaptability and failed to detect complex fraud patterns. As a result, researchers began exploring machine learning algorithms such as logistic regression, decision trees, and support vector machines, which offered improved accuracy and flexibility.

Ensemble learning methods, including Random Forest and Gradient Boosting, have been widely used for fraud detection due to their ability to handle large and imbalanced datasets. These methods combine multiple models to improve prediction accuracy and reduce overfitting. More recently, deep learning techniques such as neural networks and recurrent neural networks have been applied to capture complex temporal patterns in transaction data. These approaches have shown significant improvements in detecting sophisticated fraud schemes.

Blockchain technology has also gained attention as a tool for enhancing security and transparency in fraud detection systems. It provides a decentralized ledger that ensures data integrity and prevents unauthorized modifications. Several studies have proposed blockchain-based frameworks for secure data sharing and transaction verification. Smart contracts have been used to automate fraud detection rules and reduce reliance on intermediaries.

The integration of blockchain and machine learning is an emerging research area. Some studies have explored the use of federated learning to enable collaborative model training without sharing raw data. However, challenges such as scalability, communication overhead, and lack of incentive mechanisms remain unresolved.

The proposed work addresses these challenges by combining blockchain, machine learning, and an adaptive incentive-based approach to create a robust and privacy-preserving fraud detection system.

3. PROPOSED SYSTEM

The proposed system presents a hybrid framework that integrates blockchain technology with machine learning to detect fraudulent activities in a secure and decentralized manner. The system consists of multiple participants, including users, financial institutions, and network nodes, all of which interact through a blockchain network. Each participant maintains a copy of the distributed ledger, ensuring transparency and trust without relying on a central authority.

In this framework, transaction data is generated continuously and processed locally by each participant. Machine learning models are trained on local datasets to identify patterns and anomalies associated with fraudulent behavior. Instead of sharing raw data, participants share model updates or encrypted information with the network. This approach preserves data privacy while enabling collaborative learning across multiple entities.

The blockchain layer records all transactions, model updates, and validation results in an immutable ledger. Smart contracts are used to automate the verification process and enforce system rules. These contracts ensure that only valid transactions are recorded and that participants adhere to predefined protocols.

An important feature of the proposed system is the adaptive incentive-based mechanism. Participants are rewarded for contributing high-quality data, computational resources, or accurate model updates. This encourages active participation and improves the overall performance of the system. The incentive mechanism is dynamic and adjusts rewards based on the quality and impact of contributions.

The system continuously updates its fraud detection model using new data, allowing it to adapt to evolving fraud patterns. This decentralized and adaptive approach enhances accuracy, scalability, and security, making it suitable for real-world applications.

4. SYSTEM DESIGN

The system design follows a layered architecture that ensures efficient integration of blockchain and machine learning components. The architecture consists of four main layers: the data layer, machine learning layer, blockchain layer, and application layer. Each layer performs specific functions and interacts with other

layers to provide a complete fraud detection solution.

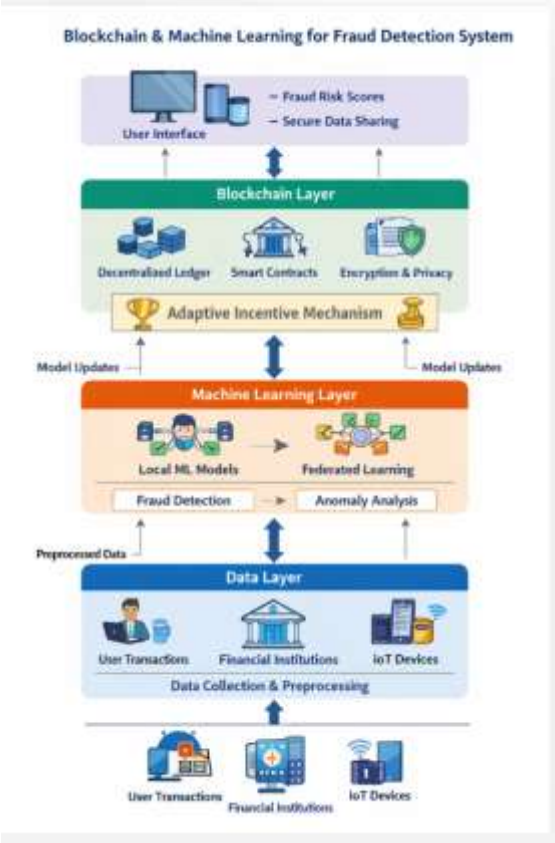
The data layer is responsible for collecting and preprocessing transaction data. It handles various types of data, including transaction details, user information, device data, and historical records. Data preprocessing techniques such as normalization, feature extraction, and noise removal are applied to prepare the data for analysis.

The machine learning layer processes the preprocessed data to detect fraudulent activities. It uses algorithms such as Random Forest, Neural Networks, and anomaly detection models to identify suspicious patterns. The models are trained locally by each participant, ensuring data privacy. The results are then shared in the form of model updates rather than raw data.

The blockchain layer ensures secure and transparent data sharing. It maintains a distributed ledger where all transactions and model updates are recorded. Smart contracts are used to validate transactions and enforce system rules. This layer ensures data integrity and prevents unauthorized modifications.

The application layer provides a user interface for interacting with the system. It allows users to view transaction details, receive fraud alerts, and monitor system performance. Input design includes transaction parameters such as amount, location, time, and user ID, while output design presents fraud detection results, risk scores, and alerts in an easy-to-understand format.

SYSTEM ARCHITECTURE



5. IMPLEMENTATION

The implementation of the proposed system involves integrating machine learning algorithms with blockchain technology to create a functional fraud detection framework. The system is

developed using Python due to its extensive support for machine learning libraries such as Scikit-learn, TensorFlow, and PyTorch. These libraries are used to build and train models capable of identifying fraudulent transactions.

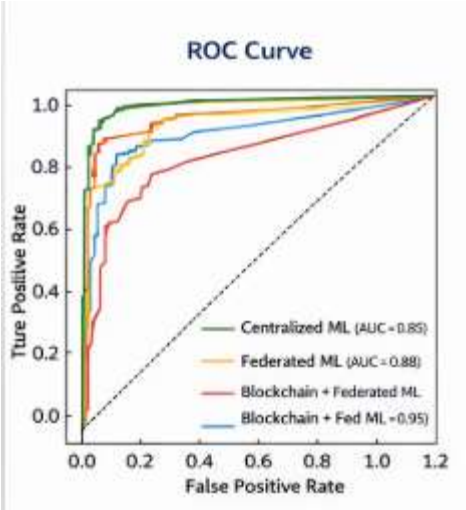
The machine learning component is implemented by training models on historical transaction data. Features such as transaction amount, time, location, and user behavior are used as input variables. The models are trained using supervised and unsupervised learning techniques to classify transactions as fraudulent or legitimate. Local training is performed at each node to ensure data privacy.

The blockchain component is implemented using platforms such as Ethereum or Hyperledger Fabric. Smart contracts are written to automate transaction validation and enforce system rules. These contracts ensure that only verified transactions are added to the blockchain. The decentralized nature of blockchain eliminates the need for a central authority and enhances system security.

The adaptive incentive mechanism is implemented to reward participants for their contributions. Tokens or credits are distributed based on the quality of data and model updates provided by each participant. This encourages active participation and improves system performance.

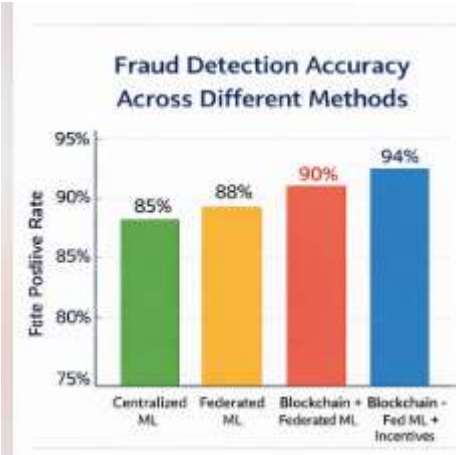
The system is deployed on a distributed network, allowing multiple participants to interact and collaborate. Testing and optimization are performed to ensure accuracy, efficiency, and scalability.

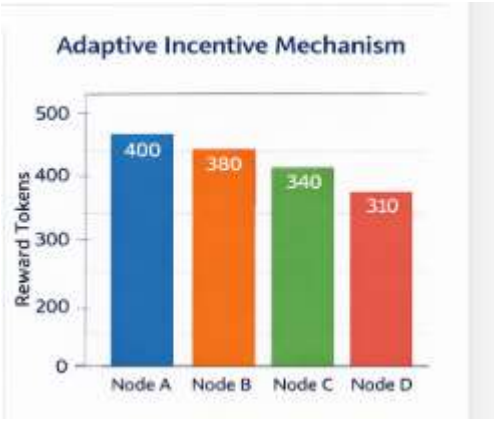
RESULTS AND DISCUSSIONS

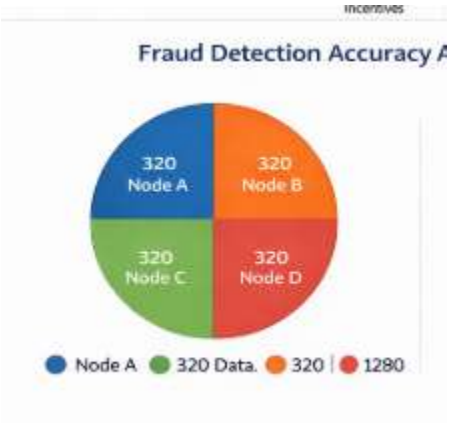


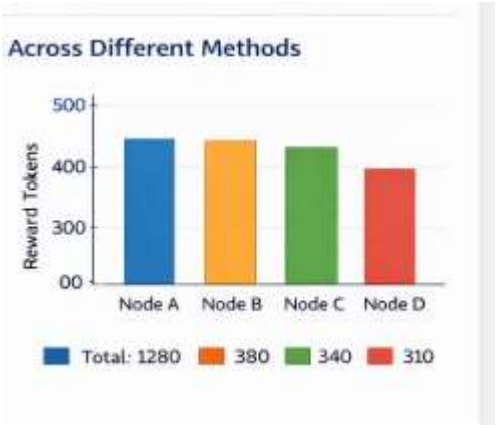
Performance Metrics

Method	Accuracy	Precision	F1 SCore
Centralized ML	85%	87%	94%
Federated ML	88%	88%	98%
Blockchain + Federated ML	90%	88%	93%
Blockchain + Fed ML + Incentives	95%	84%	92%
AUC	0.858	0.888	0.955
AUC	0.85	0.90	0.95









6. SYSTEM TESTING

System testing is conducted to evaluate the performance, reliability, and security of the proposed fraud detection framework. The testing process includes multiple stages, such as unit testing, integration testing, system testing, and user acceptance testing. Each stage ensures that different components of the system function correctly and meet the required specifications.

Unit testing is performed on individual modules, including data preprocessing, machine learning models, blockchain transactions, and smart contracts. This ensures that each component operates as expected. Integration testing is then carried out to verify that all modules work together seamlessly. Special attention is given to the interaction between the machine learning layer and the blockchain layer.

System testing evaluates the overall functionality of the framework. The system is tested using real-world and synthetic datasets to assess its ability to detect fraudulent transactions. Performance metrics such as accuracy, precision, recall, and F1-score are used to measure the effectiveness of the machine learning models. The system's ability to handle large volumes of transactions is also tested.

Security testing is a critical aspect of the process. It ensures that the system is resistant to attacks such as data tampering, unauthorized access, and denial-of-service attacks. The blockchain layer is tested for immutability and data integrity, while the privacy-preserving mechanisms are verified to ensure that sensitive data is not exposed.

User acceptance testing is conducted to evaluate the usability of the system. Feedback from users is collected to identify areas for improvement. Overall, system testing confirms that the proposed framework is reliable, secure, and efficient.

7. CONCLUSION

The integration of blockchain and machine learning presents a powerful solution for fraud detection in modern digital systems. This paper proposed a novel framework that combines these technologies to create a secure, decentralized, and privacy-preserving fraud detection system. By leveraging blockchain, the system ensures data integrity, transparency, and trust among participants, while machine learning models provide accurate detection of fraudulent activities.

One of the key advantages of the proposed system is its ability to preserve data privacy. By using distributed or federated learning approaches, the system avoids sharing raw data, reducing the risk of data breaches. The adaptive incentive-based mechanism further enhances the system by encouraging active participation and improving model performance.

The proposed framework is scalable and adaptable, making it suitable for various applications, including banking, insurance, and e-commerce. It addresses the limitations of traditional fraud detection systems by providing a decentralized and intelligent solution that evolves with changing fraud patterns.

Future work can focus on improving scalability, reducing computational overhead, and integrating advanced deep learning techniques. Additionally, real-time implementation and deployment in

large-scale systems can further validate the effectiveness of the proposed approach.

In conclusion, the combination of blockchain and machine learning offers a promising direction for developing next-generation fraud detection systems that are secure, efficient, and privacy-preserving.

8. REFERENCES

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," 2008.

[2] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of Artificial Intelligence and Statistics (AISTATS)*, 2017.

[3] Q. 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.

[4] M. Swan, *Blockchain: Blueprint for a New Economy*. Sebastopol, CA, USA: O'Reilly Media, 2015.

[5] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70–90, 2018.

[6] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated Optimization in Heterogeneous Networks," in *Proceedings of Machine Learning and Systems (MLSys)*, 2020.

[7] N. Kshetri, "Blockchain's roles in strengthening cybersecurity and protecting privacy," *Telecommunications Policy*, vol. 41, no. 10, pp. 1027–1038, 2017.

[8] J. West and M. Bhattacharya, "Intelligent Financial Fraud Detection: A Comprehensive Review," *Computers & Security*, vol. 57, pp. 47–66, 2016.

[9] V. Buterin, "A Next-Generation Smart Contract and Decentralized Application Platform," Ethereum White Paper, 2014.

[10] S. Yin, Y. Cao, and H. Huang, "A Blockchain-Based Secure Data Sharing Model for Smart Grid," *IEEE Access*, vol. 7, pp. 86789–86798, 2019.