IJSRA

International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Review Article)

Check for updates

Blockchain-based Distributed AI Models: Trust in AI model sharing

Akaash Vishal Hazarika [1, *] and Mahak Shah [2]

[1] *Department of Computer Science, North Carolina State University.*

[2] *Department of Computer Science, Columbia University.*

International Journal of Science and Research Archive, 2024, 13(02), 3493-3498

Publication history: Received on 18 November 2024; revised on 22 December 2024; accepted on 25 December 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.13.2.2598

Abstract

This paper explores the intersection of blockchain technology and dis- tributed artificial intelligence (AI) models. We analyze how blockchain can be utilized to secure AI model sharing and training processes in distributed environments, thereby enhancing trust and accountability. Key concepts such as decentralized model training, data provenance, incentivization mechanisms, and the role of smart contracts are discussed in detail. Moreover, the paper examines ethical implications and regulatory challenges inherent in this integration. Potential future research directions are outlined, emphasizing the need for enhanced scalability and interoper- ability, while promoting discussions on blockchain's implications for data ethics and AI governance.

Keywords: Blockchain; Distributed AI; Model Sharing; Trust Framework; Scalability

1. Introduction

Artificial Intelligence (AI) has significantly transformed various sectors, includ- ing finance, healthcare, logistics, and more, by enabling data-driven decision- making and automation. However, the deployment of AI models in distributed environments raises substantial concerns regarding data privacy, model integrity, and trust among participants. The traditional centralized approach to AI model training often leads to vulnerabilities, such as data breaches and biased algorithms resulting from unrepresentative training data.

1.1. Motivation

The increasing demand for collaborative AI development necessitates robust methods that can ensure data integrity, ownership, and accountability. Blockchain[1-3] technology, characterized by its decentralized and immutable nature, presents a promising solution to mitigate these challenges. By offering a secure framework for model sharing and training, blockchain can enhance trust among participant nodes and facilitate broader collaboration in AI development. Additionally, the rise of federated learning—a decentralized approach to training AI mod- els—aligns well with blockchain's strengths, as it promotes privacy-preserving collaboration across different data owners.

1.2. Objectives and Scope

This paper aims to explore the application of blockchain technology in securing AI model sharing and training processes. We focus on the following objectives:

- To analyze how blockchain can provide a secure framework for AI model sharing in distributed environments, ensuring all parties are protected from fraud and data tampering.
- To evaluate the role of smart contracts in automating and enforcing trust among participants, thus eliminating the reliance on a central authority and reducing operational bottlenecks.

[*] Corresponding author: Akaash Vishal Hazarika

- To identify challenges associated with using blockchain in AI, such as latency, transaction costs, and data normalization, and propose potential solutions to enhance system effectiveness.
- To discuss the ethical implications and regulatory considerations surround- ing this integrated approach, ensuring compliance with data protection regulations like GDPR.

1.3. Related Work

Prior research has investigated various applications of blockchain in enhanc- ing security and trust in AI systems. For example, Xu et al. [3-4] discussed the fundamental paradigms of blockchain and AI integration, while Li et al. [5] provided insights into challenges associated with federated learning methods[6] atop blockchain frameworks. Similarly, other studies have focused on the im- plications of blockchain in healthcare settings, examining how it can maintain patient privacy while allowing for collaborative AI-driven insights. Despite this progress, there remains a gap in comprehensive frameworks that integrate these technologies effectively across diverse industrial fields.

2. Background

This section provides an overview of the essential concepts underpinning blockchain technology as it relates to distributed[7-8] AI systems. Understanding these foundations enhances the application's architectural context and highlights how de- centralized technologies can reshape AI collaboration.

2.1. Blockchain Technology

2.1.1. Definition and Principles

Blockchain is a decentralized ledger technology that enables secure and transpar- ent transactions without the need for intermediaries. It operates under several principles:

- Decentralization: No single entity has control over the entire network, thereby reducing the risk of manipulation and enhancing security. Each participant (node) holds a copy of the entire blockchain, contributing to system robustness.
- Immutability[9]: Once recorded, data on the blockchain cannot be altered without consensus, ensuring the integrity of stored information. This feature is crucial in preventing fraudulent activities and ensuring trustworthy data records in model training.
- Transparency: Transactions on the blockchain are visible to all participants, promoting trust through observable accountability. This trans- parency helps in verifying data sources and model updates.
- Consensus Mechanisms: Various algorithms (such as Proof of Work and Proof of Stake) ensure agreement among network participants regard- ing the state of the ledger. These mechanisms can prevent double-spending and build consensus even in a decentralized environment.

2.1.2. Applications of Blockchain in AI

Blockchain has been proposed for several applications within AI, such as:

- Secure data exchange among disparate AI models, which facilitates collaborative learning without exposing sensitive datasets.
- Decentral training of AI algorithms to ensure data privacy while still harnessing collective insights from multiple data sources.
- Ensuring the integrity of training data and model updates, where each transaction is recorded, providing an auditable trail that enhances trust in model outputs and decision-making.
- Facilitating the tracking of model performance and user interactions, aid- ing in regulatory compliance and ethical AI practices.

2.2. Distributed AI Models

Distributed AI involves techniques where multiple agents collaborate to train models on distributed data. This paradigm addresses several challenges faced by centralized AI approaches, including:

- Data Privacy: Sensitive data can remain on local devices, only sharing model updates to ensure privacy while still benefiting from collaborative learning.
- Scalability: Distributing the training process allows for leveraging computational resources across multiple nodes, accommodating larger datasets and more complex models.
- Diversity of Data Sources: By integrating data from various partici- pants, models can become more robust and less biased, reflecting a wider range of scenarios and use cases.

2.2.1. Federated Learning

One popular approach within distributed AI is federated learning, wherein models are trained locally on nodes while only sharing gradients or model updates with a central server or between peers, thus preserving local data privacy [5]. This method:

- Mitigates the risk of data breaches by ensuring sensitive data never leaves local storage.
- Reduces the burden on network bandwidth since only aggregated updates rather than raw data are transmitted.
- Supports continuous learning in real-time, allowing AI models to adapt and improve based on recent data without major systemic overhauls.

2.2.2. Collaborative Learning

Similar to federated learning, collaborative learning involves multiple partici- pants working together to improve models by sharing insights rather than raw data. This process can be enhanced with blockchain to ensure data provenance and model versioning, providing a clear history of updates. Techniques such as differential privacy can further strengthen the security of shared insights, balancing model improvement with data confidentiality.

3. Proposed Framework

This section introduces our proposed framework for integrating blockchain with distributed AI models, designed to enhance security, trust, and accountability in collaborative AI development.

3.1. System Architecture

We provide a visual representation and detailed description of the architecture of a blockchain-based AI model-sharing system:

- Nodes: Representing participants in the network, each node can be an entity with data or computing resources, allowing for diverse and collaborative AI training.
- Smart Contracts: These automate trust enforcement and the execution of agreements made among participants, reducing the reliance on inter- mediaries and increasing efficiency in contract fulfillment.
- Decentralized Ledger: Stores model updates, transaction history, and data provenance, ensuring all participants can verify actions and outcomes.
- Off-chain Storage Solutions: Handle large datasets without bloating the blockchain, allowing for efficient data retrieval and interaction while maintaining the benefits of decentralization.

3.2. Model Sharing and Training Process

We detail how AI models can be securely shared and trained on a blockchain network through several mechanisms:

3.2.1. Data Provenance Tracking

A framework for tracking the origin and history of data used in training pro- cesses is discussed. Blockchain can support transparency and accountability by providing an immutable record of data sources, usage, and transformations throughout the AI model's lifecycle. This tracking not only enhances trust but is also crucial for compliance with data regulations.

3.2.2. Model Versioning

Methods for versioning AI models on the blockchain ensure participants can track updates and modifications. Techniques such as using digital signatures and cryptographic hashes can create unique identifiers for each model version, facilitating rollback and ensuring consistency in training.

3.2.3. Smart Contracts for Incentivization

Explore how smart contracts can be implemented to define rules for participation and reward systems for contributors to the AI training process. This includes:

- Incentive Structures for Data Providers: Rewarding data providers with tokens or credits based on data quality, quantity, and timeliness.
- Tokenomics: Designing a fair and equitable token distribution model for rewarding model trainers, contributing to sustainable development in the AI ecosystem.
- Automatic Handling of Agreements and Payouts: Ensuring that contributions and contracts are executed without delays or discrepancies, thus fostering a more streamlined process for collaboration.

4. Security and Trust

This section discusses how blockchain technology contributes to enhanced security and trust in AI model sharing, addressing issues such as data integrity and ethical implications.

4.1. Data Integrity and Provenance

Explain how blockchain enables verifiable data lineage and ensures data integrity[9-10] by storing hashes of training datasets alongside model updates. This creates an auditable record that can be accessed and verified by all participants, ensuring strict adherence to defined data standards.

4.2. Accountability and Traceability

Discuss mechanisms through which blockchain creates an audit trail for all trans- actions, empowering participants to verify actions taken in the system. A trans- parent audit trail enables stakeholders to track any changes, fostering a culture of accountability and trustworthiness in AI outputs.

4.3. Privacy Preservation Mechanisms

Investigate how advanced cryptographic techniques can be integrated into the blockchain to ensure participant privacy without compromising utility. Techniques such as homomorphic encryption and zero-knowledge proofs allow participants to validate transactions and computations without revealing underlying sensitive information.

5. Challenges and Limitations

Despite the potential benefits, there are challenges and limitations to consider when implementing blockchain in distributed AI settings:

5.1. Scalability Concerns

Examine issues related to scaling blockchain technology, including transaction speed, especially in high-volume environments. The slowness of current consen- sus mechanisms can pose barriers, requiring advancements in technology such as Layer 2 solutions or other innovations.

5.2. Energy Efficiency

Analyze the environmental impact of consensus mechanisms, particularly Proof of Work, which has garnered criticism for its energy consumption. Discuss more energy-efficient alternatives like Proof of Stake and Delegated Proof of Stake, which could mitigate these concerns while maintaining security.

5.3. Interoperability with Existing AI Frameworks

Highlight potential difficulties when integrating blockchain solutions with prevalent AI infrastructures and platforms]11-12]. This underscores the importance of establishing standardized protocols to facilitate seamless integration and ensure legacy systems can easily adopt blockchain features.

5.4. Regulatory and Ethical Considerations

Evaluate the challenges posed by varying data protection regulations (such as GDPR) and ethical concerns surrounding data usage and AI transparency. It is essential to navigate these regulations carefully to avoid non-compliance and potential legal ramifications while fostering ethical AI practices in decision- making.

6. Case Studies

Present real-world case studies or hypothetical scenarios showcasing the appli- cation of the proposed framework in diverse sectors, such as:

- Healthcare Data Sharing: Illustrate how blockchain can facilitate se- cure and compliant sharing of patient data across institutions while en- abling AI models to learn from large datasets without compromising privacy.
- Financial Sector Use Cases: Provide insights into how blockchain can be utilized for fraud detection and credit scoring, allowing institutions to collaboratively train models on historical financial transactions while ensuring data privacy.
- Supply Chain Transparency: Showcase how AI-driven predictions can improve supply chain operations by using blockchain to provide verifiable tracking of goods from producers to consumers, thus enhancing overall accountability.

Each case study should provide context, methodology, results, and a discus- sion of the implications and lessons learned, emphasizing the transformative potential of blockchain in AI applications.

7. Future Work

Suggest avenues for future research, including:

- Enhancements in scalability and efficiency of blockchain networks for AI applications, potentially through the implementation of new consensus algorithms that optimize transaction throughput while maintaining security.
- Exploration of hybrid models that incorporate cloud computing with blockchain for distributed AI training, allowing for flexibility and resource optimiza- tion.
- Investigating the legal frameworks that govern blockchain implementations in AI, focusing on how to design regulatory compliant systems while fostering innovation and development in the field.
- The impact of emerging technologies like quantum computing on blockchain and AI security, exploring how these advances could lead to new vulnera- bilities or necessitate redesigning current blockchain systems

8. Conclusion

Summarize the key findings and contributions of the paper, emphasizing that blockchain technology holds significant potential to create a secure, trustworthy environment for AI model sharing in distributed systems. Further advancements and research in this field are critical to address existing limitations and challenges, particularly in scalability, ethics, and regulatory compliance. The continued exploration of synergies between blockchain and AI could indeed lead to innovative solutions that redefine collaborative AI development across various industries.

Compliance with ethical standards

Acknowledgments

Acknowledge any contributions, collaborations, or funding sources assisting in the completion of this research. Recognition of the contributions of colleagues and collaborators can enhance the collaborative spirit of research in this field.

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] S. Nakamoto, “Bitcoin: A Peer-to-Peer Electronic Cash System,” 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] V. Buterin, “Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform,” 2013. [Online]. Available: https://ethereum. org/en/whitepaper/

[3] P. Zhang, I. Heath, and N. Foshay, “Blockchain technology in health care: A systematic review,” Health Informatics Journal, vol. 24, no. 3, pp. 1–10, 2018.

[4] J. Xu and W. Zhao, “Blockchain technology and its applications in the field of AI: A review,” Artificial Intelligence Review, vol. 52, no. 4, pp. 1073–1089, 2019.

[5] T. Li, A.K. Sahu, M. Sanjabi, and D. Phillips, “Federated Learning: Challenges, Methods, and Future Directions,” IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, 2020.

[6] A. Maraglia and B. Apolloni, “A Review on Blockchain and Its Applications in Industry 4.0,” IEEE Access, vol. 9, pp. 90704–90724, 2021.

[7] A. V. Hazarika, G. J. S. R. Ram, and E. Jain, “Performance comparison of Hadoop and Spark Engine,” in Proceedings of the 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I- SMAC), Palladam, India, 2017, pp. 671-674.

[8] A. V. Hazarika, G. J. S. R. Ram, E. Jain, D. Sushma, and Anju, “Cluster analysis of Delhi crimes using different distance metrics,” in Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 2017, pp. 565- 568.

[9] A. Chatterjee et al., “CTAF: Centralized Test Automation Framework for Multiple Remote Devices Using XMPP,” in Proceedings of the 2018 15th IEEE India Council International Conference (INDICON), IEEE, 2018.

[10] A. V. Hazarika, M. Shah, “Serverless Architectures: Implications for Distributed System Design and Implementation,” International Journal of Science and Research (IJSR), vol. 13, no. 12, pp. 1250-1253, 2024.

[11] Anju, A. V. Hazarika, “Extreme Gradient Boosting using Squared Logistics Loss function,” International Journal of Scientific Development and Research, vol. 2, no. 8, pp. 54-61, 2017

[12] M.Shah, A.V. Hazarika, “An In-Depth Analysis of Modern Caching Strategies in Distributed Systems: Implementation Patterns and Performance Implications,” International Journal of Science and Engineering Applications (IJSEA), vol. 14, no. 1, pp. 9-13, 2025.