

Survey towards Blockchain-Based Healthcare Insurance Fraud Detection

Swapnali Nagtilak¹, Ram Joshi², Nisha W Kimmatkar³, Seema Kedar⁴, Rushali Deshmukh⁵ and Prema Sahane⁶

¹⁻⁶Department of Computer Engineering, JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, 411033, India
Email: swapnali6777@gmail.com, {rbjoshi_it} @jspmrscoe.edu.in

Abstract—The escalating cost of healthcare has made cost control a critical concern for health insurance agencies. Traditional service-based reimbursement models often result in excessive and unjustifiable claims. Although the single-disease payment model has gained traction to mitigate this, it still presents a considerable risk of fraudulent activities. Insurance operations involve multiple stakeholders and are frequently hindered by outdated, opaque claim processing mechanisms. Blockchain technology, with its immutable ledger and decentralized validation, offers a promising solution to address these inefficiencies and fraud vulnerabilities. This paper proposes a consortium blockchain-based framework enhanced with machine learning to detect fraudulent insurance claims automatically. The system includes an explainable AI component that assesses the legitimacy of medical codes based on patient symptoms. Moreover, the use of blockchain ensures secure, transparent, and auditable storage of healthcare records, thereby assisting auditors and reinforcing trust in the claim settlement process.

Index Terms— Healthcare, Machine Learning, Insurance Fraud AI, Blockchain.

I. INTRODUCTION

Health insurance (HI) serves as a financial agreement wherein insurers compensate subscribers for healthcare expenditures arising from medical conditions, injuries, or accidents. As healthcare costs surge globally, HI has become essential for individuals, not only for medical coverage but also for tax-saving incentives. Historically, claim processing in health insurance involved manual, paper-based procedures that required in-person visits to insurance offices, making the process time-consuming and prone to errors or manipulation.

Digital transformation has significantly streamlined the health insurance claim (HIC) process by offering automated services that minimize human intervention and enhance operational efficiency. Some benefits include easier access for policyholders, reduced dependency on intermediaries, faster claim verification, and the ability to detect anomalies using artificial intelligence (AI). However, digital HIC systems are not without challenges. Issues such as data validation, talent shortages in algorithmic modeling, and the evolving nature of insurance fraud remain prevalent.

Notably, fraud in HI has reached alarming levels, with fraudulent claims accounting for a significant percentage of public funds in schemes like Medicare. The absence of effective penalties and detection mechanisms exacerbates the issue. Additionally, many insurers still operate on legacy systems that lack integration with modern technologies, increasing operational costs and security risks. The growing reliance on digital platforms also raises serious concerns regarding user privacy and data misuse, especially when subscriber data is shared

without explicit consent. Cyberattacks such as ransomware, phishing, and denial-of-service have further exposed vulnerabilities in the current health insurance ecosystem. Encryption methods do help in this[32] but , there is a need to develop fraud-resistant, secure, and transparent claim processing systems.

II. RELATED WORK

A. Securing Electronic Health Records Using Blockchain and IPFS

This work introduces a decentralized electronic health record (EHR) system that leverages blockchain and the Inter Planetary File System (IPFS) to enhance data privacy and security. The proposed model integrates cryptographic controls and role-based access, allowing patients to grant permission to authorized participants verified via digital certificates. By categorizing data and assigning access roles, the system ensures immutability, transparency, and accountability without third-party intervention. A prototype was implemented to assess cost efficiency, and optimization strategies were suggested for system improvement [1].

B. Fundamentals of Blockchain and Its Role in Smart Healthcare

This study offers an introductory exploration of blockchain technology, tracing its origin back to timestamping solutions in the 1990s and its evolution through the launch of Bitcoin by Satoshi Nakamoto in 2009. The paper outlines blockchain's core properties—such as decentralization, distributed ledgers, and open access—and discusses its expanding utility across industries. In healthcare, these characteristics make blockchain particularly suitable for secure and verifiable data management, offering potential applications in EHRs, insurance processing, and medical supply chains[2].

C. Predicting Insurance Fraud Using Naïve Bayes and K-NN

This research evaluates the effectiveness of two machine learning algorithms—Naïve Bayes and K-Nearest Neighbors (K-NN)—in predicting fraudulent insurance claims. The case study focuses on low-premium policies taken out for elderly clients, which are later exploited through coverage enhancements and false claims. A dataset of 14,778 claim records was used, with key features such as claim date, policy value, and region. The models were trained and tested using an 80:20 split, and their performance was measured using confusion matrices and AUC values. The study emphasizes early-stage fraud detection at policy issuance [3].

D. Blockchain and AI-Driven Framework for Healthcare Insurance Fraud Detection

Addressing the growing concern of fraud in health insurance, this paper presents a comprehensive survey of AI- and blockchain-integrated solutions for detecting anomalies in medical claims. It identifies critical challenges in data sensitivity, security, and volume. A hybrid framework is proposed, combining blockchain's immutability and traceability with the predictive power of AI to automate fraud identification. The system's effectiveness is illustrated through a case study, and a taxonomy of common fraud scenarios and attack vectors is also provided for clarity.[4]

E. Review of Insurance Claim Fraud Detection Systems

This review highlights the persistent issue of insurance fraud, which continues to inflict financial losses on the industry globally. With India's insurance market poised for significant growth, the need for robust fraud prevention systems becomes crucial. The study introduces a blockchain-based framework for secure, peer-to-peer communication among stakeholders and proposes the use of smart contracts to streamline operations. It categorizes different types of insurance fraud—health, vehicle, and life—and maps them to corresponding machine learning detection techniques. Future research directions are discussed, focusing on system integration and real-time analytics [5].

F. Blockchain Model for Reducing Risk and Fraud in the Insurance Sector

This research focuses on building a blockchain-driven model to enhance security and minimize fraud in the insurance industry. Traditional claim processing mechanisms are identified as major risk factors due to their centralized and manual nature. Through qualitative observation of existing workflows and business practices, the authors identify key vulnerabilities and propose blockchain as a solution for fraud-proof, transparent claim recording. A prototype model is developed, emphasizing the reduction of risk through immutable, tamper-resistant ledgers, and a decentralized verification process[6].

TABLE I. LITERATURE REVIEW

Sr.No.	Paper Title	Technology Used	Application
1	Issues and Effectiveness of Blockchain Technology on Digital Voting.[7]	Blockchain Technology	Proposes a backup and recovery mechanism for health blockchain keys using body sensor networks, showcasing how similar lightweight frameworks can be adapted for healthcare fraud prevention.
2	Electronic voting machine based on Block chain technology and Aadhar verification[8]	Blockchain Technology	Highlights smart contracts and security modules that ensure secure patient data management—applicable in healthcare insurance for access control and claim verification.
3	E-Voting with Blockchain: An E-Voting Protocol with Decentralisation and Voter Privacy[9]	Blockchain Technology	Introduces a consortium blockchain structure for auditability and secure data sharing; a model transferable to healthcare data auditing and fraud tracking.
4	Design of Distributed Voting Systems[10]	Blockchain Technology	Describes a blockchain and cloud- integrated system for secure personal health data exchange; applicable in distributed healthcare insurance networks.
5	Finding the needle: A risk-based ranking of product listings at online auction sites for non- delivery fraud prediction[19]	ML Techniques	Discusses ranking methods for fraud-prone listings using imbalanced data and cost- sensitive learning; methodology is extendable to insurance fraud detection scenarios
6	Credit card fraud detection using machine learning techniques: A comparative analysis[20]	ML Techniques	Explores classification techniques under changing fraud patterns and skewed datasets; provides foundational insight for developing adaptive fraud detection models.
7	Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms [21]	ML Techniques	This paper compares multiple models for detecting corporate financial misreporting and serves as a guide for evaluating ML models in healthcare insurance data.
8	Credit Card Fraud Detection using Machine Learning Algorithms[22]	ML Techniques	Proposes a real-time transaction behaviour analysis method to flag anomalies, which can be repurposed for streaming healthcare claims.
9	Fraud Detection using Machine Learning and Deep Learning[23]	Machine Learning and Deep Learning	Combines deep learning with conventional techniques for transaction fraud detection; relevant for complex healthcare datasets requiring high accuracy.
10	Credit Card Fraud Detection using Machine Learning Algorithms[25]	ML Algorithms	Evaluates Random Forest and AdaBoost to determine the most accurate algorithm; a model selection strategy that can inform healthcare fraud detection systems.

III. METHODOLOGY

This review adopts a structured thematic approach to analyze recent literature on blockchain-based healthcare fraud detection. Peer-reviewed journal articles and conference papers published between 2020 and 2023 were considered. The selection was based on their relevance to the integration of blockchain with machine learning for health insurance applications, particularly in fraud prevention and claim validation [25].

The methodology involved three core stages:

1. **Literature Collection:** Sources were gathered using digital libraries like IEEE Xplore, SpringerLink, and ScienceDirect, with specific emphasis on systems utilizing smart contracts, AI, and decentralized architectures.
2. **Thematic Classification:** Identified papers were categorized based on the technologies used (e.g., blockchain type, ML algorithms), target application (e.g., fraud detection, claim automation), and evaluation methods (e.g., simulation, prototype, theoretical analysis).
3. **Comparative Evaluation:** Key features, limitations, and performance insights were extracted to assess the potential of various frameworks. Emphasis was placed on architectural design, data privacy assurance, fraud detection mechanisms, and scalability[26].

This methodology ensured that the review remains focused on the intersection of emerging technologies and real-world healthcare insurance fraud detection challenges.

IV. PROPOSED FRAMEWORK

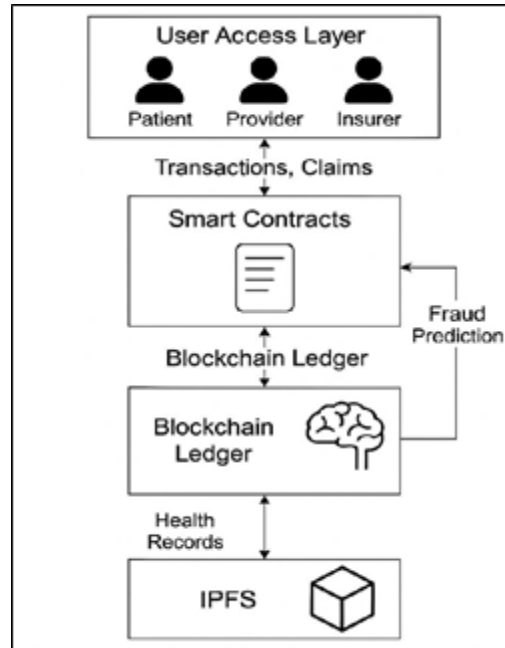


Fig1: System Architecture

The proposed framework introduces an integrated system architecture aimed at mitigating fraud in healthcare insurance through the convergence of blockchain, smart contracts, and machine learning.

1. **Blockchain Ledger:** Acts as a tamper-proof, decentralized database to log all medical and insurance transactions. It enhances transparency and supports audit trails.
2. **Smart Contracts:** Self-executing protocols are embedded into the blockchain to automate the verification of claim conditions. These contracts enable trustless enforcement of policy rules and instant flagging of anomalies.
3. **AI Engine:** Employs machine learning models to analyze historical claim patterns and detect suspicious behaviors in real time. To improve accuracy and adapt to evolving fraud tactics, the system learns from the new data[27].
4. **IPFS (InterPlanetary File System):** Supports the decentralized storage of large-scale medical records and diagnostic reports, ensuring data integrity and easy retrieval without centralized dependency.
5. **User Access Layer:** Provides interfaces for various stakeholders—including patients, insurance firms, and healthcare providers—to access, upload, and validate data securely with role-based permissions.

Together, these components form a secure, scalable, intelligent system for proactive fraud detection and transparent healthcare insurance operations.

IV. CONCLUSIONS

The implementation of blockchain technology in healthcare insurance presents a transformative opportunity to enhance data security, transparency, and fraud prevention. Its decentralized nature, combined with smart contracts and immutable records, provides a robust foundation for tamper-proof claim processing[28]. This review underscores that while blockchain alone improves traceability, its true potential lies in conjunction with intelligent models such as machine learning. The integration of AI facilitates real-time anomaly detection, supports audit decisions, and reduces manual oversight. The system overcomes limitations of traditional centralized architectures by decentralizing storage through IPFS[29]. The proposed framework offers a holistic approach to modernizing insurance processes and curbing fraudulent activities. While challenges related to regulatory compliance, interoperability, and system scalability remain, future research can address these by exploring hybrid blockchain models, federated learning, and privacy-preserving techniques.

REFERENCES

- [1] M. Ali, S. Ahmed, M. I. Hossain, A. B. M. Alim Al Islam, and J. Noor, "Electronic health record's security and access control using blockchain and IPFS," in **Proc. 7th Int. Congr. Inf. Commun. Technol.**, Singapore: Springer, 2023, pp. 493–505.
- [2] S. Gupta, H. K. Sharma, and M. Kapoor, "Introduction to blockchain and its application in smart healthcare system," in **Blockchain for Secure Healthcare Using Internet of Medical Things (IoMT)**, Cham: Springer, 2023, pp. 55–65.
- [3] G. Triyono and D. Ginting, "Comparative analysis performance of Naïve Bayes and K-NN using confusion matrix and AUC to predict insurance fraud," **J. Media Informatika Budidarma**, vol. 6, no. 4, pp. 2293–2300, 2022.
- [4] K. Kapadiya et al., "Blockchain and AI-empowered healthcare insurance fraud detection: An analysis, architecture, and future prospects," **IEEE Access**, vol. 10, pp. 79606–79627, 2022.
- [5] S. Vyas and S. Serasiya, "Fraud detection in insurance claim system: A review," in **Proc. 2nd Int. Conf. Artif. Intell. Smart Energy (ICAIS)**, 2022, pp. 922–927, doi: 10.1109/ICAIS53314.2022.9742984.
- [6] S. E. Hiererra et al., "Blockchain technology for fraud detection and risk prevention in insurance industry," **Unpublished manuscript**.
- [7] A. Gupta, J. Patel, M. Gupta, and H. Gupta, "Issues and effectiveness of blockchain technology on digital voting," **Int. J. Eng. Manuf. Sci.**, vol. 7, no. 1, 2017.
- [8] A. Navya, R. Roopini, and A. S. SaiNiranjan, "Electronic voting machine based on blockchain technology and Aadhar verification," **Int. J. Adv. Res. Ideas Innov. Technol.**, vol. 4, no. 2, 2018.
- [9] F. S. Hardwick, R. N. Akram, and K. Markantonakis, "E-voting with blockchain: An e-voting protocol with decentralisation and voter privacy," **arXiv preprint arXiv:1805.10258**, 2018.
- [10] C. Meter, "Design of distributed voting systems," **arXiv preprint arXiv:1702.02566**, 2017.
- [11] S. Panja and B. K. Roy, "A secure end-to-end verifiable e-voting system using zero knowledge-based blockchain," **Unpublished**.
- [12] M. A. Makary and M. Daniel, "Medical error—the third leading cause of death in the US," **BMJ: British Med. J. (Online)**, vol. 353, 2016.
- [13] P. T. S. Liu, "Medical record system using blockchain, big data and tokenization," in **Int. Conf. Inf. Commun. Security**, Springer, 2016, pp. 254–261.
- [14] A. Decker and R. Wattenhofer, "Information propagation in the Bitcoin network," in **Proc. 13th IEEE Int. Conf. Peer-to-Peer Comput.**, 2013, pp. 1–10.
- [15] T. Neudecker, P. Andelfinger, and H. Hartenstein, "Timing analysis for inferring the topology of the Bitcoin peer-to-peer network," in **IEEE Int. Conf. UIC/ATC/ScalCom/CBDCCom/IoP/SmartWorld**, 2016, pp. 358–367.
- [16] D. Zhang, "Resilience enhancement of container-based cloud load balancing service," **PeerJ Preprints**, 2018.
- [17] Y. Kou, C. T. Lu, and S. Sirwongwattana, "Survey of fraud detection techniques," in **Proc. IEEE Int. Conf. Networking, Sensing and Control**, 2004.
- [18] A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," **J. Network Comput. Appl.**, Elsevier, 2016.
- [19] V. Almendra, "Finding the needle: A risk-based ranking of product listings at online auction sites for non-delivery fraud prediction," **Expert Syst. Appl.**, 2013. [Online]. Available: <https://colab.ws/articles/10.1016%2Fj.eswa.2013.02.027>
- [20] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in **Proc. Int. Conf. Comput., Netw. Informatics (ICCN)**, IEEE, 2017.
- [21] J. Perols, "Financial statement fraud detection: An analysis of statistical and machine learning algorithms," **Auditing: J. Pract. Theory**, AAA, 2011.
- [22] V. N. Dornadula and S. Geetha, "Credit card fraud detection using machine learning algorithms," **Procedia Comput. Sci.**, Elsevier, 2019.
- [23] P. Raghavan and N. El Gayar, "Fraud detection using machine learning and deep learning," in **Proc. Int. Conf. Intell. Comput. Control Syst. (ICICCS)**, IEEE, 2019.
- [24] A. Thennakoon, C. Bhagyan, and I. Perera, "Real-time credit card fraud detection using machine learning," in **Int. Conf. Cloud Comput. Intell. Syst.**, IEEE, 2019.
- [25] R. Sailusha, V. Gnaneswar, and R. Ramesh, "Credit card fraud detection using machine learning," in **Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)**, IEEE, 2020.
- [26] Anonymous, "Sequence classification for credit-card fraud detection," **Expert Syst. Appl.**, 2018.
- [27] Kimmattkar, Nisha Vishnupant, and B. Vijaya Babu. "Novel approach for emotion detection and stabilizing mental state by using machine learning techniques." *Computers* 10.3 (2021): 37.
- [28] Kimmattkar, Nisha Vishnupant, and B. Vijaya Babu. "Human emotion detection with electroencephalography signals and accuracy analysis using feature fusion techniques and a multimodal approach for multiclass classification." *Engineering, Technology & Applied Science Research* 12.4 (2022): 9012-9017.
- [29] P. Sahane, A. Obla and P. Bhargale, "Student Achievements Portfolio Management using Blockchain," 2023 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2023, pp. 1787-1792, doi: 10.1109/ICICT57646.2023.10134076.
- [30] R. Deshmukh, T. Lahamge, I. Phadatare, D. Shinde and R. Manhas, "Real-Time Marathi Sign Language Recognition using Deep Learning Techniques," 2024 3rd International Conference on Automation, Computing and Renewable

Systems (ICACRS), Pudukkottai, India, 2024, pp. 1326-1331, doi: 10.1109/ICACRS62842.2024.10841608.

- [31] Kedar, Seema, Saurabh Sutar, and Himanshu Prasad. "Smart Analyzer: Assisting College Management through Machine Learning and Data Analysis." *Turkish Journal of Computer and Mathematics Education* 12.1S (2021): 137-145.]
- [32] Manjrekar, A.A., Bhosale, A.H., Khandagale, H.P., Nikam, S.S. (2025). Data Storage with Attribute-Based Encryption and De-duplication Filter on Hybrid Cloud. In: Tuba, M., Akashe, S., Joshi, A. (eds) *ICT Systems and Sustainability. ICT4SD 2024. Lecture Notes in Networks and Systems*, vol 1159. Springer, Singapore. https://doi.org/10.1007/978-981-97-8526-1_42.