Journal of Autonomous Intelligence (2024) Volume 7 Issue 4
doi: 10.32629/jai.v7i4.1274

ORIGINAL RESEARCH ARTICLE

Machine learning for effective EHR management in blockchain-cloud integration

Birendra Kumar Saraswat[1,*], Aditya Saxena[1], P. C. Vashist[2]

[1] *GLA University, Chaumuhan, Mathura, Uttar Pradesh 281406, India*

[2] *G L Bajaj Institute of Technology and Management, Greater Noida, Uttar Pradesh 201306, India*

*** Corresponding author:** Birendra Kumar Saraswat, saraswatbirendra@gmail.com

ABSTRACT

Machine learning (ML) techniques have gained prominence in effectively managing Electronic Health Record (EHR) systems within the context of blockchain-cloud integration. This study presents a hybrid Machine Learning approach that combines logistic regression (LR) and random forest (RF) techniques for EHR management, leveraging the data stored in a blockchain-cloud integrated system. The tamper-resistant nature of blockchain ensures the authenticity and security of the stored patient information, serving as a reliable source for learning. The proposed LR+RF model is evaluated against other algorithms, considering various performance metrics. The analysis reveals that the LR+RF model achieves an impressive accuracy rate of 98.37%, indicating its efficacy in accurately classifying EHR data and facilitating effective management. Furthermore, the study compares the performance of blockchain-cloud-based decentralized storage with blockchain-based storage and peer-to-peer storage in terms of latency and throughput. The results demonstrate that the blockchain-cloud integrated decentralized storage surpasses other storage methods, achieving an average throughput of 6.8 units and a latency of 4.7 units. These findings highlight the potential of the proposed LR+RF model for EHR management within a blockchain-cloud integrated environment. The use of blockchain as a secure storage environment ensures the integrity of patient information, while Machine Learning techniques enhance the accuracy of classification.

Keywords: healthcare; electronic health record; machine learning; blockchain; cloud computing; security; privacy

ARTICLE INFO

Received: 21 September 2023
Accepted: 17 November 2023
Available online: 2 February 2024

1. Introduction

EHRs have been breakthroughs in the medical industry because they digitize patient health information, make it widely available, and improve the effectiveness of healthcare delivery. Traditional methods of organizing and analyzing EHRs confront considerable hurdles as the amount and complexity of healthcare data continue to expand[1]. In recent years, an effective approach has emerged to overcome these difficulties and improve the efficacy of EHR administration by integrating Machine Learning techniques, blockchain technology, and cloud computing. Machine learning is a subfield of AI that enables computers to learn from data and make accurate choices and forecasts on their own. EHR data may be processed, analyzed, and interpreted with the use of Machine Learning algorithms to discover useful insights, spot patterns, and bolster clinical decision-making. Machine Learning methods can be helpful in many areas of medicine, including diagnosis, therapy prediction, risk assessment, and individualized care. However, a safe system for storing data, exchanging it, and working together is essential for the widespread use of Machine Learning in

healthcare[2]. This is where blockchain technology and cloud computing may truly emerge.

Blockchain technology's distributed and unchangeable ledger protects EHR data without compromising privacy or usability[3]. Data may be sent in an encrypted and unalterable format without compromising privacy or security. The cloud, on the other hand, provides scalable and inexpensive options for archiving, analyzing, and disseminating enormous EHR datasets across a wide range of healthcare providers and stakeholders. The difficulties of EHR management can be overcome through the synergistic combination of Machine Learning, blockchain, and cloud computing[4]. Data privacy, security, and interoperability are all ensured by blockchain technology[5], allowing healthcare businesses to use Machine Learning algorithms to derive actionable insights from EHR data. The storage and processing capabilities made possible by the cloud allow Machine Learning models to access and analyze vast volumes of data.

1.1. Benefits of blockchain-cloud integration in EHR management

Blockchain and cloud computing, when brought together, offer an effective solution to the problems associated with EHR management. Data privacy, integrity, and compatibility are at stake in EHRs because they contain private and personal patient health information that must be protected throughout storage, retrieval, and transmission. Blockchain's distributed ledger technology[6] and cloud computing's scalability and adaptability can help healthcare firms build an effective environment for EHR administration[7]. The distributed, immutable ledger provided by blockchain technology ensures the integrity and authenticity of all transactions. **Figure 1** given below illustrates the architecture of a Blockchain-cloud-based EHR system.

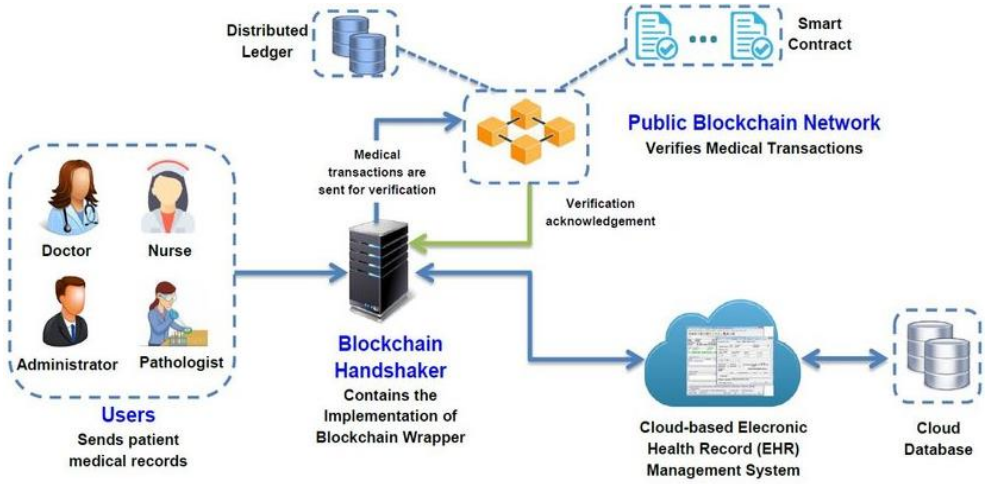

Figure 1. Blockchain-cloud-based EHR management system[8].

There are several advantages to using blockchain for EHR administration. To begin with, it safeguards information by generating an immutable log of all transactions, which is extremely difficult to modify in any way without detection. This function is essential for ensuring the integrity of patient's medical records. Second, blockchain improves the safety and privacy of user data. Centralized databases used by most conventional EHR systems are a common weakness that must be addressed. Blockchain technology, on the other hand, uses cryptographic techniques to encrypt and secure data, making it so that only authorized users may access and read certain pieces of electronic health records[9]. Patient information is more secure because of this decentralized approach to data storage and access management, which reduces the likelihood of data breaches and single points of failure. Blockchain also facilitates safe and reliable information exchange between various players in the healthcare industry. Due to the heterogeneity of healthcare IT and the absence of defined data formats, interoperability has been a persistent problem in EHR administration. Blockchain technology allows for the safe and uniform storage and transfer of electronic health records (EHRs) between institutions. Because of this interoperability, medical professionals may examine a more complete picture of a patient's health and

make more educated decisions and better coordinate their treatment for the individual[10]. **Figure 2** illustrates a blockchain based system in the healthcare sector.

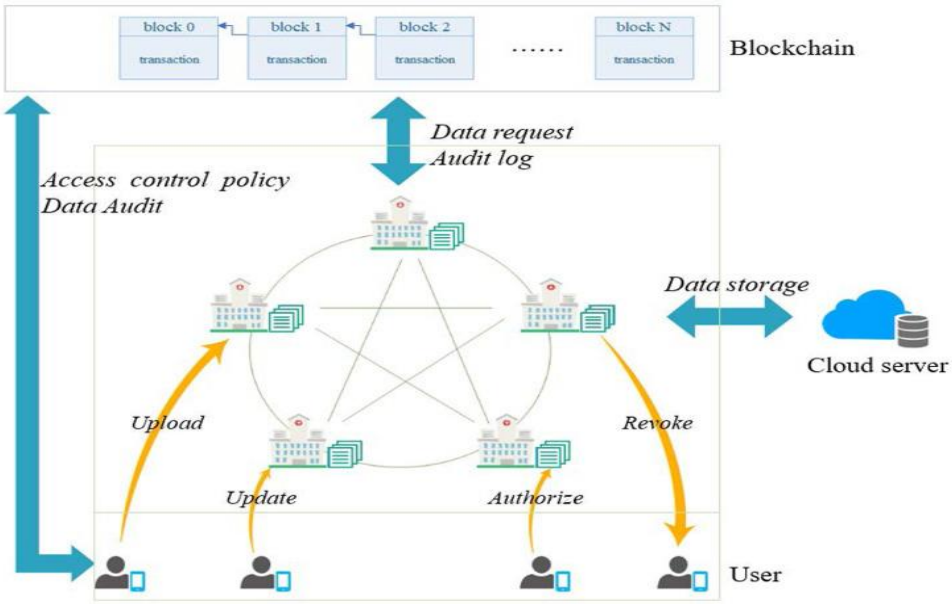

Figure 2. Blockchain-based systems in the healthcare sector[11].

The cloud enables the storage, processing, and access to massive amounts of EHR data, making it an ideal partner for blockchain technology. As the quantity of data created by EHRs continues to grow, cloud platforms provide a scalable and flexible solution, allowing healthcare businesses to do so without having to invest in expensive hardware infrastructure. Advanced data analysis, predictive modeling, and decision assistance are all made possible by the cloud's ability to easily integrate with Machine Learning and analytics technologies. A secure and scalable EHR management platform may be established for healthcare companies through the combination of blockchain and cloud computing. Blockchain technology ensures data integrity, transparency, and immutability, while cloud computing guarantees data availability, scalability, and efficiency. The entire potential of EHR data for research, population health management, and customized medication may now be tapped by healthcare practitioners because of this integration.

The combination of blockchain and cloud computing in EHR administration represents a game-changing strategy for the safekeeping, dissemination, and analysis of patient health records. Healthcare firms may improve data privacy, integrity, and interoperability, as well as take advantage of cloud computing's scale and flexibility, by merging the two technologies. Healthcare delivery, patient outcomes, and industry innovation might all benefit from the blockchain, and cloud being successfully integrated into EHR administration.

1.2. Role of machine learning in EHR management

The efficiency of EHR management may be greatly improved with the use of Machine Learning. Making use of Machine Learning algorithms, healthcare organizations may mine EHR data for actionable intelligence, which can then be used to enhance clinical decision-making and evaluate the EHR system's overall performance. Predictive modeling and risk assessment are where Machine Learning shines most in EHR administration. Large amounts of historical EHR data may be analyzed by Machine Learning algorithms, revealing patterns and connections that human analysts would miss. In order to anticipate patient outcomes, identify health hazards, and spot illness early warning signals, these algorithms may learn from data and construct predictive models. With this foresight, doctors may take preventative measures that benefit patients and save costs.

Decision-support systems for healthcare also benefit greatly from the use of Machine Learning[12]. Personalized therapy suggestions may be generated by Machine Learning algorithms by evaluating EHR data, considering patient characteristics, medical history, and reactions to previous interventions. These algorithms can help doctors choose the best course of therapy for their patients depending on the available data. Furthermore, Machine Learning may automate the extraction of significant clinical details from unorganized EHR data allowing for faster and more precise diagnosis and treatment planning.

The efficiency of the EHR system may also be evaluated using Machine Learning techniques. Machine Learning may analyze user activity, system logs, and past data to pinpoint bottlenecks in EHR processes and suggest solutions. These algorithms may be programmed to automatically correct mistakes, eliminate wasteful repetition, and suggest improvements to workflow and user experience. Healthcare businesses may improve patient care and fuel innovation by tapping into the full potential of EHR data using Machine Learning algorithms. Medical research, population health management, and clinical practice might all benefit greatly from the incorporation of Machine Learning into EHR systems[13].

There is enormous potential for improving patient care, healthcare delivery, and EHR administration through the combination of Machine Learning, blockchain, and cloud computing. Healthcare companies may enhance patient outcomes and population health management by utilizing these technologies to their fullest extent and discovering the full potential of EHR data. Hence, this study aims to examine how Machine Learning techniques might be used to blockchain-cloud integration for efficient EHR administration. In this study, Machine Learning algorithms are utilized to improve clinical decision-making and patient care outcomes by gathering useful information from EHR data. This study also looks into the scalability and cost-effectiveness of cloud computing for enabling ML-based EHR management, and the significance of blockchain technology in protecting the confidentiality, authenticity, and integrity of electronic health record data.

2. Literature review

This section presents an overview of the relevant work conducted by various authors in the field of EHR Management in Blockchain-Cloud integration using various approaches.

Zhang et al.[14] suggested a blockchain-based privacy-preserving e-health system to solve the security issues pursuing the current cloud-assisted EHRs. This study discussed the dangers of EHRs being tampered with or leaked by unscrupulous medical professionals or cloud storage service providers. The authors offer pairing-based cryptography to create immutable records incorporated into blockchain transactions, therefore protecting the privacy of electronic health information. The electronic health records of the patients are protected from unauthorized changes and may be verified with this method. The study also covers the development of safe payment protocols utilizing blockchain-based smart contracts for trustworthy payments between patients and hospitals for diagnostic and storage services. Validation via security analysis and performance assessment demonstrates the efficacy and low computational cost of the proposed approach.

Ismail et al.[15] explored healthcare blockchain-cloud integration (BcC), or the use of blockchain technology with cloud computing. The study utilized the scalability and effectiveness of cloud computing in combination with the decentralized nature of blockchain to address security and privacy issues. In the study, the authors surveyed all aspects of BcC integration in healthcare, including the various architectures, apps, and development tools currently in use. challenges, solutions, and plans for the future of the field were also discussed. The study's findings can aid the healthcare sector in improving patient care through the use of new data management systems.

Velmurugadass et al.[16] developed a new method of criminal investigation that makes use of blockchain technology. Mobile nodes, an open-flow switch, blockchain-based controllers, a cloud server, an Authentication Server (AS), and investigators are all parts of the framework's Cloud-based Software Defined

Network (SDN). For information safety, the system makes use of cryptographic algorithms and cryptographic hash functions based on the Elliptic Curve Integrated Encryption Scheme (ECIES). Based on a Logical Graph of Evidence (LGoE), the investigators carry out several tasks, such as identification, evidence collecting, analysis, and report preparation. Response time, accuracy, throughput, and security characteristics were all significantly enhanced in experimental findings. Criminal investigations and evidence management might benefit from the integration of blockchain, SDN, and encryption methods, as demonstrated by this study.

Benil and Jasper[17] presented a novel approach to deal with EHR security concerns called Elliptical Curve Certificateless Aggregate Cryptography Signature (EC-ACS). The system safeguards private medical records and prevents unauthorized access by utilizing approved blockchain technology. Medical records are encrypted using Elliptic Curve Cryptography (ECC), and digital signatures are generated using the Certificateless Aggregate Signature (CAS) approach, both of which help to make cloud storage and sharing possible. The suggested technique safeguards the cloud-based healthcare system by enforcing confidentiality and preventing illegal access. Integrating blockchain technology further ensures the integrity, traceability, and secure cloud storage of medical records.

Shi et al.[11] performed a comprehensive literature assessment of blockchain options for EHR systems with an emphasis on security and privacy. The study set out to investigate blockchain's potential utility in EHR systems and to spot gaps and openings in the field. The writers emphasized the rising interest in blockchain's revolutionary potential in the healthcare industry. They did, however, note that several obstacles remain in the way of the complete integration of blockchain technology with traditional EHR systems. Some of these difficulties were reviewed, and potential topics for further study were highlighted, including the Internet of Things (IoT), big data, ML, and edge computing. The aging society may greatly benefit from the creation of next-generation EHR systems, which the authors of this study intend to facilitate.

Bhattacharya et al.[18] proposed Blockchain-Based Deep Learning as a Service (BinDaaS) as a framework to solve the issues of confidentiality, security, as well as data integrity in EHRs. The system combines blockchain technology with deep learning algorithms to enable the safe transfer of EHR data between different medical institutions. In the first stage, lattice-based cryptography is presented as an authentication and signature technique that can withstand collusion attempts across healthcare authorities. To forecast future illnesses based on patient indications and attributes, the second step entails employing Deep Learning as a Service (DaaS) on archived EHR information. Accuracy, end-to-end latency, mining time, and computation and transmission expenses are only a few of the metrics used to gauge the success of the suggested system. The results show that BinDaaS performs better than competing solutions across all of these measures, making it the best option for managing and predicting EHRs.

Guo et al.[19] suggested a multi-authority attribute-based signing technique to guarantee the integrity of blockchain-stored EHRs. Patients can now recommend attribute-based messaging without disclosing any more personal information using this system. Since the blockchain is a decentralized ledger, using several authorities eliminates the requirement for a single point of failure. The protocol protects itself against collusion attacks by having its authorities share private pseudorandom function (PRF) seeds. When compared to existing approaches, the suggested attribute-based signature system proved to be both secure and efficient. This method, when combined with blockchain technology, gives patients more control over their medical records and allows for the safe, distributed administration of electronic health records.

Al Omar et al.[20] focused on the growing interest of cybercriminals regarding medical data and recommended a blockchain-based, patient-centric approach to managing healthcare records. MediBchain was a solution that used a decentralized network of peers to protect user privacy and data integrity. In order to maintain the privacy of their patients' information and attain pseudonymity, the researchers used Elliptic Curve Cryptography (ECC) for encryption and cryptographic functions. The review highlighted the benefits of the

proposed platform while also presenting a privacy-preserving approach for healthcare data. The goal of this study was to create a decentralized system that would improve the patient experience on the web while protecting their privacy.

Hasanova et al.[21] presented a machine learning-based algorithm called Sine Cosine Weighted K-Nearest Neighbour (SCA-WKNN) is proposed for the early prediction of heart disease. The algorithm utilizes data stored in the tamper-resistant blockchain, ensuring data authenticity and secure storage for patient information. The performance of SCA-WKNN is compared to other algorithms using metrics such as accuracy, precision, recall, F-score, and root mean square error. The results show that SCA-WKNN achieves higher accuracy compared to W K-NN and KNN, with an improvement of 4.59% and 15.61%, respectively. Furthermore, the study compares blockchain-based storage with peer-to-peer storage in terms of latency and throughput, finding that decentralized blockchain storage has a maximum throughput 25.03% higher than peer-to-peer storage. This research highlights the potential of IoT integration and blockchain technology in facilitating early detection and secure management of heart disease.

3. Problem formulation and research objectives

EHRs are stored in many healthcare systems. It's challenging to verify their legitimacy, privacy, and security. Blockchain technology makes EHR management decentralized and secure. However, blockchain technology's implementation in healthcare is still in its infancy, and an EHR framework that integrates blockchain, cloud computing, and ML to better EHR administration is needed. This study leverages blockchain and cloud computing to establish a secure, effective, and scalable EHR platform. The proposed framework would employ ML classification techniques to evaluate the EHR system. This cutting-edge technology would create an EHR system that secures, protects, and verifies patient data. Healthcare providers may manage and exchange patient data more easily with adaptable, cost-effective technology. Machine Learning approaches could enable the system to analyse data and improve patient care. In general, combining blockchain, cloud computing, and Machine Learning in an EHR framework could change healthcare data management. EHR management and patient outcomes become secure, effective, and scalable with this solution.

The research objectives of this study include:

- To create a blockchain-based EHR system that is secure and scalable enough to handle the huge amount of data generated by the healthcare industry.
- To integrate cloud computing to provide a cost-effective and scalable EHR management platform.
- To develop classification algorithms based on Machine Learning that can accurately assess the efficacy of the EHR system.
- To confirm data privacy and security by implementing the right encryption and access control in the EHR system.

4. Dataset description

MIMIC-III is a widely used healthcare dataset available on Kaggle. It stands for "Medical Information Mart for Intensive Care III". The MIMIC-III dataset contains de-identified electronic health records of over 40,000 critical care patients. It includes a vast array of data, such as demographic information, vital signs, laboratory measurements, medications, procedures, and more. MIMIC-IIIc refers to the aggregated version of the dataset, which combines multiple subsets into a single cohesive dataset. Researchers and data scientists often utilize MIMIC-IIIc to explore and develop healthcare-related models and applications[22]. The MIMIC-III dataset includes:

- Patient demographics (age, gender, ethnicity, admission/discharge dates)
- Clinical data (vital signs, lab measurements, medications, procedures, diagnoses)
- Notes and reports (progress notes, discharge summaries, radiology reports, nursing documentation)

- Procedures and interventions (surgeries, intubation, ventilation, dialysis, medication administration)
- Severity scores (SAPS, SOFA, MPM) to assess illness severity and predict outcomes.
- ICU-specific data (admission/discharge times, length of stay, ICU care team details)
- Data linkage for comprehensive analyses across different aspects of patient care.

5. Research methodology

The healthcare industry is currently exploring the integration of blockchain technology, cloud storage, and Machine Learning classifiers in an EHR infrastructure. This study aims to leverage blockchain's immutability and cloud computing's scalability to securely store and process healthcare data. A research methodology ensures the credibility and legitimacy of the study's findings, ensuring a methodical and orderly approach. The study involves explaining the dataset and methods used to gather and analyse necessary data for achieving objectives. By utilizing cutting-edge technologies, this study aims to create a secure and scalable EHR infrastructure, enabling efficient and accurate management and exchange of patient data by healthcare providers. The resulting system would improve patient outcomes while maintaining utmost confidentiality and privacy of patient information.

5.1. Proposed architecture

This section outlines the proposed methodology in several steps. It commences with patient registration to create an EHR at a medical center, employing HMAC for patient-doctor authentication. The EHR data is then generated, encrypted using ECC, and securely stored on the Blockchain. Data sharing is facilitated via a cloud server using a pre-shared secret key. Upon request from another medical center, the encrypted EHR data is decrypted with ECC and provided, completing the process. The EHR data log is logged in a database, subsequently divided into train and test data for Machine Learning classification, utilizing classifiers like Random Forest and Logistic Regression for evaluation. The proposed methodology's architecture is depicted in **Figure 3**.

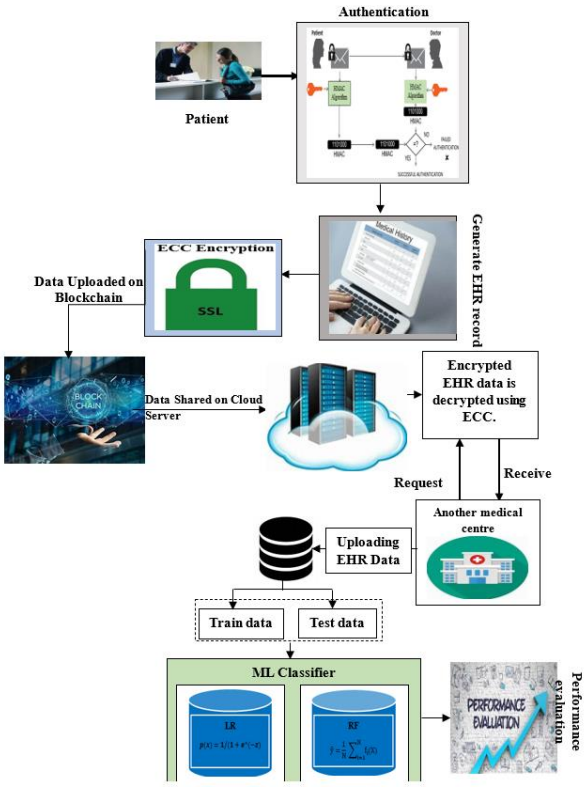

Figure 3. Proposed architecture.

Proposed algorithm

The author represents some of the individual steps using mathematical notations:

Step 1: Patient registration:

Let P be the set of patients, where P = {p1, p2, ..., pn}.

When a new patient registers, add them to the set P:

P' = P U {pn+1}

Step 2: HMAC authentication:

The authentication code can be generated using the HMAC function as follows:

authentication_code = H (key, patient_info)

Step 3: Encryption using ECC:

Let E (m, k) be the ECC encryption function, where 'm' is the message and 'k' is the public key.

The encrypted EHR data can be represented as:

encrypted_data = E (EHR_data, public_key)

Step 4: Blockchain upload:

Let B be the Blockchain and T be the transaction that contains the encrypted EHR data along with the previous and current block hashes.

The upload of encrypted EHR data on Blockchain can be represented as:

T = {encrypted_data, previous_block_hash, current_block_hash}

B = B U {T}

Step 5: Verification of key condition:

Let K be the pre-shared secret key and received_key be the key received from the other medical center requesting the data.

The key verification condition can be represented as:

IF (received_key == K)

THEN receive_data()

ELSE end_process()

Step 6: ML classification:

Let X be the input data, Y be the output labels, and f be the classifier function.

The application of ML classifiers (Random Forest and Logistic Regression) for classification can be represented as:

Y_predicted = f(X_train)

accuracy = accuracy_score(Y_test, Y_predicted)

6. Result and discussion

The study aimed to evaluate the efficiency of the proposed data processing mechanism using the Blockchain-Cloud integrated system. By storing patient data in this system, secure access to sensitive health data is ensured. Healthcare professionals can access this data to predict diseases. The following section presents the results of the performance achieved by the ML algorithm in accurately assessing the efficacy of the EHR system and the performance achieved to store patient data on a blockchain-cloud integrated system.

6.1. ML algorithms in assessing EHR system efficacy

The study introduced LR+RF algorithms for processing patient data, known for their effectiveness in handling medical data and achieving accurate predictions. The performance of the LR+RF algorithm was compared with the SCA-WKNN algorithm and other traditional ML methods. The comparative analysis

presented in **Table 1** provides a comprehensive assessment of the algorithms' performance in accurately processing patient data and predicting diseases.

Table 1. Comparative analysis of proposed ML techniques with existing techniques.

Techniques	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	RMSE
SCA_WKNN[21]	92.13	88.21	93.27	90.66	0.1115
KNN[21]	81.49	69.7	62	65.62	0. 25
SVM[21]	82.7	79	71.03	74.80	0.41
RF[21]	89.14	81	81.04	81.02	0.18
MLP[21]	90.14	87.13	85.13	86.11	0.123
Proposed LR+RF[21]	98.37	94.83	99.64	97.17	0.1031

6.1.1. Based on accuracy

The graph illustrated in **Figure 4** indicates that the proposed LR+RF technique achieved the highest accuracy among all the methods evaluated in this study, with an accuracy rate of 98.37%. This demonstrates the effectiveness of the proposed LR+RF technique for the given task, surpassing the accuracy rates of other techniques. The SCA-WKNN technique also performed well with an accuracy of 92.13%, showing its potential for accurate predictions. The K-NN, SVM, RF, and MLP techniques achieved accuracy rates of 81.49%, 82.7%, 89.14%, and 90.14%, respectively. While these methods demonstrated decent performance, the proposed LR+RF technique outperformed them by a significant margin.

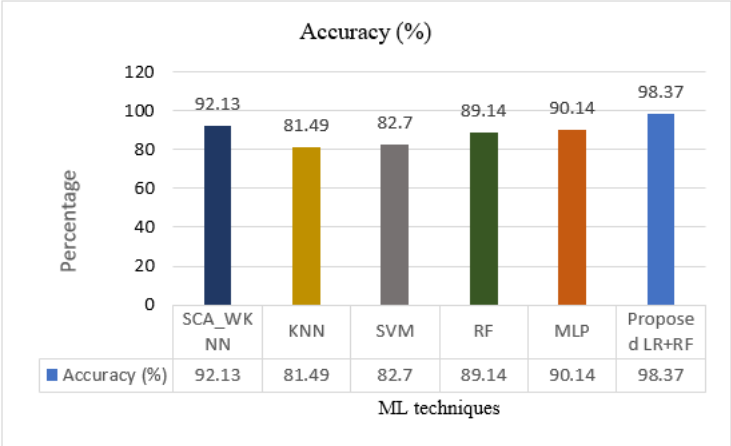

Figure 4. Accuracy of various machine learning techniques.

6.1.2. Based on precision

It could be seen from the graph given in **Figure 5** that SCA_WKNN demonstrated a higher precision value at 88.21%. It showcases the effectiveness of the SCA_WKNN algorithm in accurately classifying electronic health records. While KNN achieved a precision of 69.7%, it may not be the most suitable choice for this task. SVM and RF algorithms demonstrated better performance with precisions of 79% and 81%, respectively. The MLP algorithm exhibited a precision of 87.13%, indicating its capability to handle the complexities of EHR data in the context of Blockchain-Cloud integration. However, the proposed hybrid model, LR+RF, outperformed all other techniques, achieving an impressive precision of 94.83%. This model combines the strengths of LR and RF to deliver superior performance in accurately managing electronic health records.

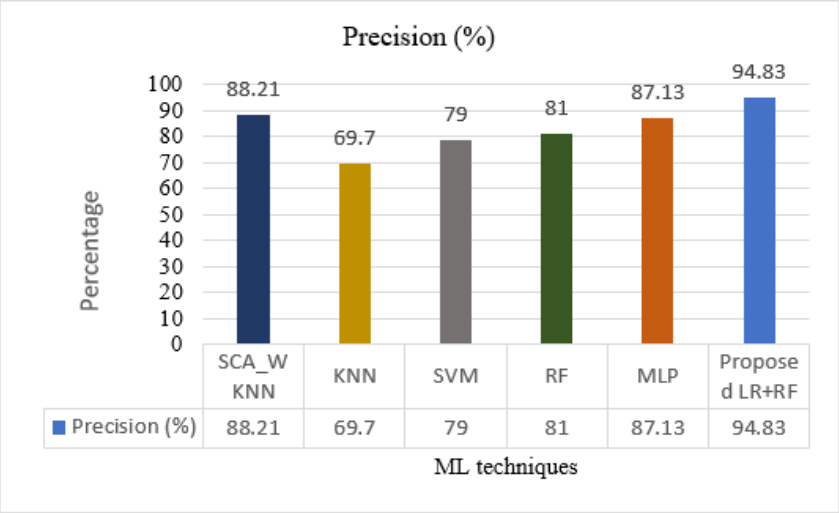

Figure 5. Precision of various Machine Learning Techniques.

6.1.3. Based on recall

It could be observed from the graph given in **Figure 6** that SCA_WKNN achieved a higher recall rate at 93.27%. This demonstrates the effectiveness of the SCA_WKNN algorithm in accurately identifying relevant information within the EHR system. However, KNN exhibited a lower recall rate of 62%, suggesting that it may not be the most suitable choice for this task. SVM and RF algorithms demonstrated better performance, achieving recall rates of 71.03% and 81.04%, respectively. The MLP algorithm exhibited a recall rate of 85.13%, showcasing its capability to capture important patterns and information from the EHR system. Notably, the proposed LR+RF model, outperformed all other techniques, achieving an exceptional recall rate of 99.64%. This indicates the superior ability of the LR+RF model to accurately assess the efficacy of the EHR system in the context of Blockchain-Cloud integration. The proposed LR+RF hybrid model, with its exceptional recall rate, offers promising potential for the comprehensive and accurate evaluation of the EHR system's performance.

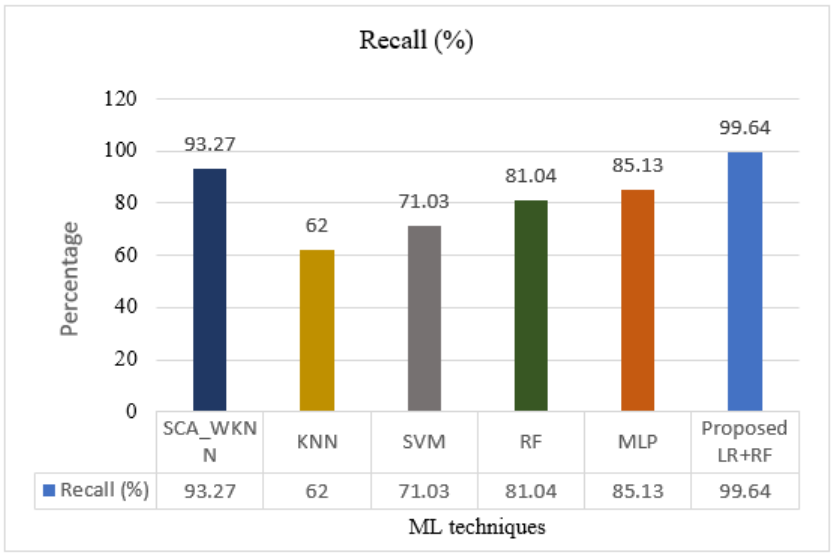

Figure 6. Recall of various machine learning techniques.

6.1.4. Based on F1-score

The graph given in **Figure 7** illustrates that the SCA_WKNN achieved the highest F1-score of 90.66%. This indicates that SCA_WKNN is effective in assessing the efficacy of the EHR system within the

Blockchain-Cloud Integration environment. On the other hand, KNN exhibited a lower F1 score of 65.62%, suggesting that it may not be the most suitable algorithm for accurately evaluating the performance of the EHR system in this context. SVM and RF algorithms performed better, with F1 scores of 74.8% and 81.02% respectively, indicating their competence in assessing the efficacy of the EHR system. The MLP algorithm showcased a respectable F1-score of 86.11%, implying its capability to handle the intricacies of the EHR data within the Blockchain-Cloud Integration environment. However, the proposed hybrid model, LR+RF, outperformed all other techniques with an outstanding F1 score of 97.17%. This hybrid model, which combines the strengths of Logistic Regression (LR) and Random Forests (RF), demonstrates exceptional performance in accurately assessing the efficacy of the EHR system.

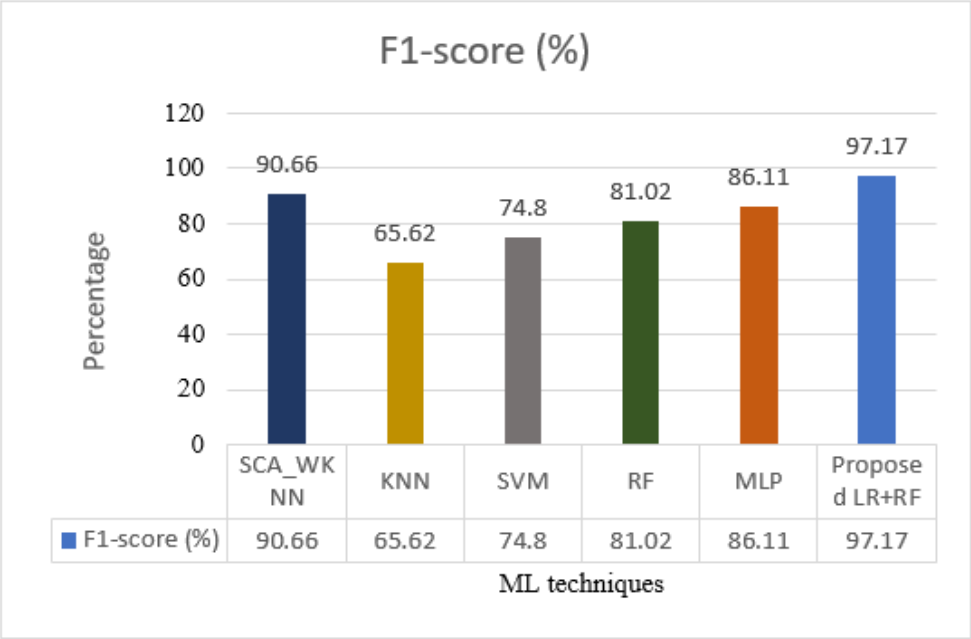

Figure 7. F1-Score of various Machine Learning Techniques.

6.1.5. Based on root mean square error (RMSE)

The graph given in **Figure 8** depicts that the proposed LR+RF model achieved the lowest RMSE value of 0.1031, indicating its superior performance in accurately assessing the efficacy of the EHR system. The combination of Logistic Regression (LR) and Random Forests (RF) proved effective in capturing the complexities of the data and generating more precise predictions. While SCA_WKNN and MLP also performed well with RMSE values of 0.115 and 0.123, respectively, KNN and SVM exhibited higher RMSE values of 0.25 and 0.41, suggesting less accurate assessments in this context. The RF algorithm demonstrated a competitive performance with an RMSE of 0.18, indicating its effectiveness in evaluating the EHR system's efficacy.

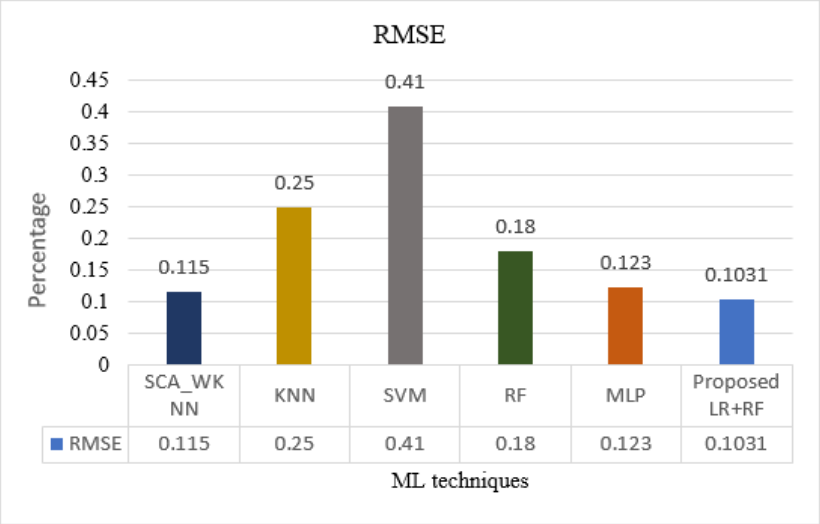

Figure 8. RMSE of various machine learning techniques.

6.2. Performance evaluation of blockchain-cloud integrated storage

Further, the study utilized decentralized blockchain-cloud integrated approaches to storing patient data and the results observed are as follows:

6.2.1. Based on latency

Figure 9 shows that as the number of users in the network grows, so does the delay. While the blockchain-cloud integrated solution, blockchain-based decentralized storage, and peer-to-peer storage all have their advantages, they all share a similar strength: optimum latency. The blockchain-cloud integrated method achieves a latency of 0.085 milliseconds (ms) when there are 200 patients, whereas blockchain-based storage and peer-to-peer storage reach latencies of 0.158 ms and 0.17 s, respectively. The blockchain-cloud integrated method achieves a latency of 0.37 ms when the number of patients exceeds 400, whereas blockchain-based storage and peer-to-peer storage both record latencies of 0.455 ms and 0.514 ms, respectively. When compared to other storage solutions, blockchain-cloud-integrated decentralized storage achieves an average latency of 0.28 ms and consistently displays excellent latency over the whole range of patient counts.

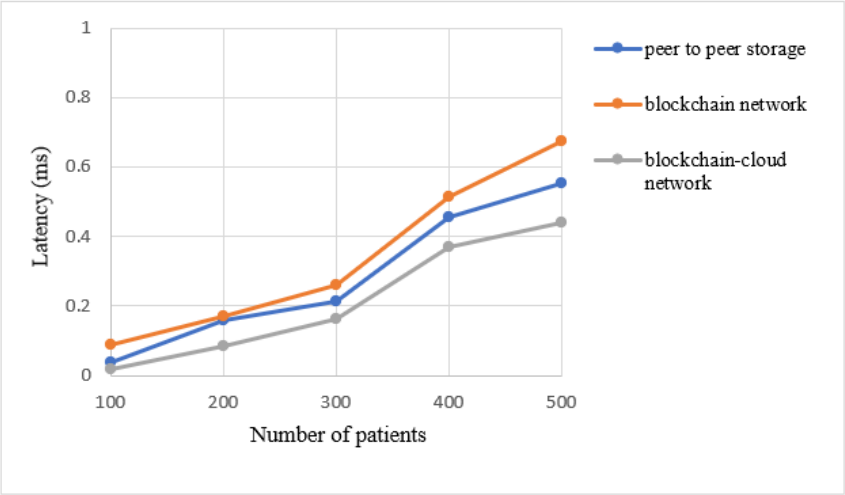

Figure 9. Latency comparison.

6.2.2. Based on throughput

The next step involves testing blockchain-cloud integrated decentralized storage, block-chain based storage, and peer-to-peer storage to determine throughput. Compared to both peer-to-peer and blockchain-

based storage, blockchain-cloud integrated storage achieves the maximum throughput as seen in **Figure 10**. At 300 patients, the throughput of blockchain-cloud integrated storage is 0.48 megabits per second (Mbps), while that of blockchain-based storage is 0.54 Mbps and that of peer-to-peer storage is 0.48 Mbps. In terms of throughput, blockchain-cloud integrated storage achieves an average throughput of 0.65 Mbps and outperforms blockchain-based and peer-to-peer storage throughout the board, for any given number of patients.

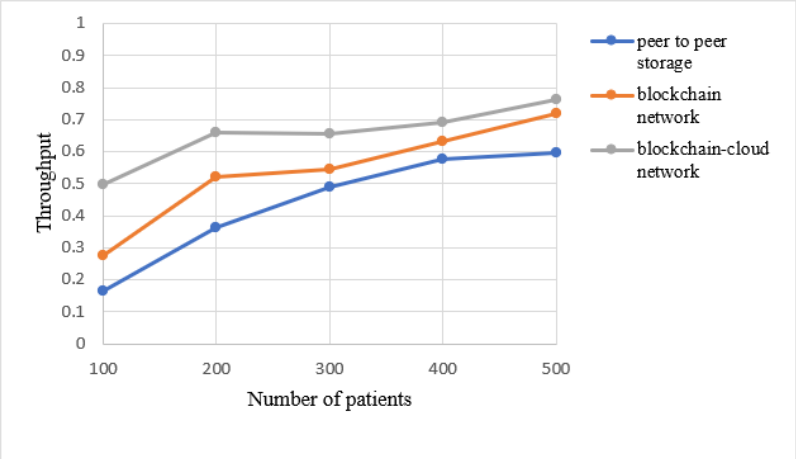

Figure 10. Throughput comparison.

7. Conclusion and future scope

In this study, the application of Machine Learning techniques for effective EHR management within a blockchain-cloud integrated environment is investigated. The results demonstrate the potential of Machine Learning in improving the accuracy and efficiency of EHR management, while also addressing the challenges of data security and privacy in a decentralized setting. Based on the results obtained from the evaluation of different Machine Learning techniques for effective Electronic Health Record (EHR) management in a blockchain-cloud integration environment, it can be concluded that the proposed LR+RF model outperformed other techniques in terms of accuracy, precision, recall, F1-score, and RMSE. With an accuracy of 98.37% and high values for precision, recall, and F1-score, the LR+RF model demonstrates its effectiveness in EHR management. Furthermore, when comparing storage solutions, the blockchain-cloud integrated decentralized storage consistently exhibited optimal latency performance, maintaining an average latency of 0.28 ms across various patient counts. This highlights the efficiency of the blockchain-cloud integration approach in handling increasing numbers of patients while maintaining low latency. In terms of throughput, the blockchain-cloud integrated storage showcased superior performance with an average throughput of 0.65 Mbps. It consistently outperformed blockchain-based and peer-to-peer storage solutions for all patient counts, indicating its ability to handle data transfer at a higher rate.

Future studies in Machine Learning for EHR management in a blockchain-cloud integrated environment should prioritize improved algorithms, enhanced data security, interoperability, real-world implementation, and ethical considerations. Advanced algorithms, encryption techniques, and privacy-preserving methods should be developed to enhance accuracy and protect sensitive data. Establishing interoperability standards and conducting real-world studies will assess practicality and scalability. Ethical frameworks must guide responsible Machine Learning usage. Addressing these areas will unlock the full potential of Machine Learning, revolutionizing EHR management and advancing healthcare systems.

Author contributions

Conceptualization, BKS; methodology, BKS; software, AS and PCV; validation, BKS, AS and PCV; resources, BKS; data curation, BKS; writing—original draft preparation, BKS, AS and PCV; writing—review

and editing, AS and PCV; visualization, AS and PCV; supervision, BKS, AS and PCV; project administration. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Tang F, Ma S, Xiang Y, et al. An Efficient Authentication Scheme for Blockchain-Based Electronic Health Records. IEEE Access. 2019, 7: 41678-41689. doi: 10.1109/access.2019.2904300
2. Gianfrancesco MA, Tamang S, Yazdany J, et al. Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. JAMA Internal Medicine. 2018, 178(11): 1544. doi: 10.1001/jamainternmed.2018.3763
3. Shahnaz A, Qamar U, Khalid A. Using Blockchain for Electronic Health Records. IEEE Access. 2019, 7: 147782-147795. doi: 10.1109/access.2019.2946373
4. Premarathne U, Abuadbba A, Alabdulatif A, et al. Hybrid Cryptographic Access Control for Cloud-Based EHR Systems. IEEE Cloud Computing. 2016, 3(4): 58-64. doi: 10.1109/mcc.2016.76
5. Pandey P, Litoriya R. Securing and authenticating healthcare records through blockchain technology. Cryptologia. 2020, 44(4): 341-356. doi: 10.1080/01611194.2019.1706060
6. Yang G, Li C. A design of blockchain-based architecture for the security of electronic health record (EHR) systems. In: 2018 IEEE International conference on cloud computing technology and science (CloudCom). pp. 261-265. IEEE, 2018.
7. Ganiga R, Pai RM, M. M. MP, Sinha RK. Security framework for cloud based electronic health record (EHR) system. International Journal of Electrical and Computer Engineering (IJECE). 2020, 10(1): 455. doi: 10.11591/ijece.v10i1.pp455-466
8. Rahman MS, Khalil I, Mahawaga Arachchige PC, et al. A Novel Architecture for Tamper Proof Electronic Health Record Management System using Blockchain Wrapper. Proceedings of the 2019 ACM International Symposium on Blockchain and Secure Critical Infrastructure. 2019. doi: 10.1145/3327960.3332392
9. Tanwar S, Parekh K, Evans R. Blockchain-based electronic healthcare record system for healthcare 4.0 applications. Journal of Information Security and Applications. 2020, 50: 102407. doi: 10.1016/j.jisa.2019.102407
10. Mayer AH, da Costa CA, Righi R da R. Electronic health records in a Blockchain: A systematic review. Health Informatics Journal. 2019, 26(2): 1273-1288. doi: 10.1177/1460458219866350
11. Shi S, He D, Li L, et al. Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. Computers & Security. 2020, 97: 101966. doi: 10.1016/j.cose.2020.101966
12. Gultepe E, Green JP, Nguyen H, et al. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. Journal of the American Medical Informatics Association. 2014, 21(2): 315-325. doi: 10.1136/amiajnl-2013-001815
13. Wong J, Murray Horwitz M, Zhou L, et al. Using Machine Learning to Identify Health Outcomes from Electronic Health Record Data. Current Epidemiology Reports. 2018, 5(4): 331-342. doi: 10.1007/s40471-018-0165-9
14. Zhang G, Yang Z, Liu W. Blockchain-based privacy preserving e-health system for healthcare data in cloud. Computer Networks. 2022, 203: 108586. doi: 10.1016/j.comnet.2021.108586
15. Ismail L, Materwala H, Hennebelle A. A Scoping Review of Integrated Blockchain-Cloud (BcC) Architecture for Healthcare: Applications, Challenges and Solutions. Sensors. 2021, 21(11): 3753. doi: 10.3390/s21113753
16. Velmurugadass P, Dhanasekaran S, Shasi Anand S, et al. Enhancing Blockchain security in cloud computing with IoT environment using ECIES and cryptography hash algorithm. Materials Today: Proceedings. 2021, 37: 2653-2659. doi: 10.1016/j.matpr.2020.08.519
17. Benil T, Jasper J. Cloud based security on outsourcing using blockchain in E-health systems. Computer Networks. 2020, 178: 107344. doi: 10.1016/j.comnet.2020.107344
18. Bhattacharya P, Tanwar S, Bodkhe U, et al. BinDaaS: Blockchain-Based Deep-Learning as-a-Service in Healthcare 4.0 Applications. IEEE Transactions on Network Science and Engineering. 2021, 8(2): 1242-1255. doi: 10.1109/tnse.2019.2961932
19. Guo R, Shi H, Zhao Q, et al. Secure Attribute-Based Signature Scheme With Multiple Authorities for Blockchain in Electronic Health Records Systems. IEEE Access. 2018, 6: 11676-11686. doi: 10.1109/access.2018.2801266
20. Omar AA, Bhuiyan MZA, Basu A, et al. Privacy-friendly platform for healthcare data in cloud based on blockchain environment. Future Generation Computer Systems. 2019, 95: 511-521. doi: 10.1016/j.future.2018.12.044
21. Hasanova H, Tufail M, Baek UJ, et al. A novel blockchain-enabled heart disease prediction mechanism using machine learning. Computers and Electrical Engineering. 2022, 101: 108086. doi: 10.1016/j.compeleceng.2022.108086

22. MIMIC3c aggregated data. Available online: https://www.kaggle.com/datasets/drscarlat/mimic3c (accessed on 7 December 2023).