International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

PRIVACY-PRESERVING EHR ANALYTICS USING FEDERATED LEARNING AND BLOCKCHAIN

[1] U. ARAVIND, [2] KOLAGATLA PRANEETH KUMAR REDDY, [3] SHAIK MASTAN, [4] MUKTHIPUDI RITHIK KUMAR, [5] SHAIK HUSSAIN SHARIF, [6] BOMMIREDDY HEMANTH KUMAR REDDY

[1] ASST., PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES – DEVARAJUGATTU, MARKAPUR

[2,3,4,5,6] STUDENT, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES – DEVARAJUGATTU, MARKAPUR

ABSTRACT

The rapid digitization of healthcare systems has led to the widespread adoption of Electronic Health Records (EHR), raising critical concerns regarding data privacy, security, and interoperability. Traditional centralized machine learning approaches for healthcare analytics often require direct access to sensitive patient data, increasing the risk of data breaches and violating regulatory compliance. To address these challenges, this paper proposes a blockchain-integrated framework for secure federated learning of encrypted EHR data using homomorphic encryption.

The proposed system combines the decentralized nature of blockchain technology with the privacy-preserving capabilities of federated learning. In this framework, healthcare institutions collaboratively train machine learning models without sharing raw patient data. Instead, local models are trained on-site using encrypted EHR data, and only model updates are transmitted. Homomorphic encryption ensures that computations can be performed directly on encrypted data, eliminating the need for decryption and thereby enhancing confidentiality.

Keywords: Blockchain, Federated Learning, Electronic Health Records (EHR), Homomorphic Encryption, Data Privacy, Secure Machine Learning, Distributed Learning, Smart Contracts, Healthcare Data Security, Privacy-Preserving AI, Cryptography, Decentralized Systems

I. INTRODUCTION

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

The healthcare industry is undergoing a significant digital transformation driven by the adoption of **Electronic Health Records (EHR)**, advanced analytics, and artificial intelligence. These technologies enable improved diagnosis, personalized treatment, and efficient clinical decision-making. However, the increasing reliance on digital health data has introduced serious concerns related to data privacy, security, and regulatory compliance. Sensitive patient information stored in EHR systems is highly vulnerable to breaches, unauthorized access, and misuse, making secure data handling a critical challenge in modern healthcare ecosystems.

Traditional machine learning approaches in healthcare typically rely on centralized data collection, where patient data from multiple institutions is aggregated into a single repository for model training. While effective in generating accurate predictive models, this approach poses significant risks, including data leakage, violation of privacy laws such as HIPAA, and the creation of single points of failure. Moreover, healthcare organizations are often reluctant to share data due to legal, ethical, and competitive concerns, leading to data silos that limit the effectiveness of machine learning systems.

II. LITERATURE REVIEW

The integration of **blockchain**, **federated learning**, and **homomorphic encryption** for secure healthcare data analysis has gained significant attention in recent years. Researchers have explored various approaches to address privacy, security, and scalability challenges associated with **Electronic Health Records (EHR)**.

Several studies have focused on the application of **federated learning** in healthcare systems. McMahan et al. introduced the concept of federated learning, enabling decentralized model training without sharing raw data. Their work demonstrated how distributed clients can collaboratively build machine learning models while preserving data privacy. Later research extended this concept to healthcare applications, showing its effectiveness in disease prediction and medical imaging analysis. However, these approaches highlighted vulnerabilities such as information leakage through model updates and reliance on centralized aggregation servers.

To overcome these limitations, researchers have explored **homomorphic encryption** as a privacy-preserving mechanism. Gentry's pioneering work on fully homomorphic encryption (FHE) enabled

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

computations on encrypted data without decryption. Subsequent studies applied homomorphic encryption in healthcare analytics to secure sensitive EHR data during processing. Although highly secure, these methods often suffer from high computational overhead, making them challenging to deploy in real-time healthcare environments.

EXISTING SYSTEM

In current healthcare data analytics, most systems rely on **centralized architectures** for storing, processing, and analyzing **Electronic Health Records (EHR)**. In this traditional approach, data from multiple hospitals, clinics, and healthcare providers is collected and stored in a central database or cloud server. Machine learning models are then trained using this aggregated dataset to perform tasks such as disease prediction, patient risk assessment, and clinical decision support.

One widely used method involves transferring sensitive patient data to centralized servers where data preprocessing, feature extraction, and model training are performed. Although this approach allows for high model accuracy due to access to large-scale datasets, it introduces significant privacy and security concerns. Sensitive medical information becomes vulnerable to cyberattacks, unauthorized access, and data breaches. Moreover, strict regulations and compliance requirements (such as healthcare data protection laws) limit the ability of institutions to share patient data freely.

To address privacy concerns, some existing systems have adopted **basic encryption techniques** during data storage and transmission. However, these methods typically require decryption before computation, exposing the data during processing. This creates a potential risk window where sensitive information can be compromised.

In recent years, **federated learning** has been introduced as an alternative approach. In this system, data remains within local institutions, and only model updates are shared with a central server for aggregation. While federated learning improves privacy by avoiding raw data sharing, it still depends on a centralized aggregator, which can become a single point of failure. Additionally, model updates can sometimes leak sensitive information through inference attacks, reducing the overall security of the system.

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

PROPOSED SYSTEM

The proposed system introduces a **blockchain-integrated framework for secure federated learning of encrypted Electronic Health Records (EHR)** using **homomorphic encryption**. This system is designed to overcome the limitations of existing centralized and partially secure approaches by ensuring **end-to-end data privacy, decentralized trust, and secure collaborative model training** across multiple healthcare institutions.

In this framework, each participating hospital or healthcare provider acts as a **local client** that retains its patient data within its own infrastructure. Instead of sharing raw EHR data, local machine learning models are trained on-site using encrypted datasets. **Homomorphic encryption** is applied to the data, allowing computations to be performed directly on encrypted information without revealing the actual content. This ensures that sensitive patient data remains protected at all stages of processing.

The system utilizes **federated learning** to enable collaboration among multiple institutions. After local training, only encrypted model updates (such as weights or gradients) are shared with the network. Unlike traditional federated learning, which depends on a centralized aggregator, the proposed system replaces this component with a **blockchain network**. This eliminates the single point of failure and enhances trust among participants.

A **blockchain layer** is integrated to manage and record all transactions related to model updates. Each update is verified and stored in an immutable distributed ledger, ensuring transparency, traceability, and resistance to tampering. **Smart contracts** are deployed to automate key processes such as participant authentication, validation of model updates, aggregation rules, and incentive mechanisms. This guarantees that only authorized entities contribute to the learning process and that malicious activities are prevented.

METHODOLOGY

The proposed system follows a structured methodology that integrates federated learning, homomorphic encryption, and blockchain technology to ensure secure and privacy-preserving analysis of Electronic Health Records (EHR). Initially, participating healthcare institutions register on the blockchain network and are authenticated through cryptographic identities managed by smart contracts. Each institution maintains its local dataset and performs preprocessing tasks such as data cleaning, normalization, and feature extraction within its secure

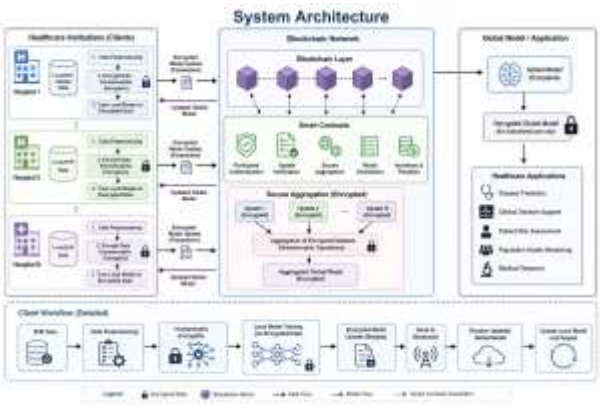
International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

environment. The EHR data is then encrypted using homomorphic encryption techniques, allowing computations to be executed without exposing sensitive information.

Once the data is prepared, local machine learning models are trained independently at each institution using the encrypted data. After training, only the encrypted model parameters (such as gradients or weights) are generated and shared with the blockchain network instead of raw data. These updates are broadcast to the decentralized network, where smart contracts verify their authenticity and integrity. The blockchain acts as a secure ledger, recording each transaction and ensuring transparency and immutability.

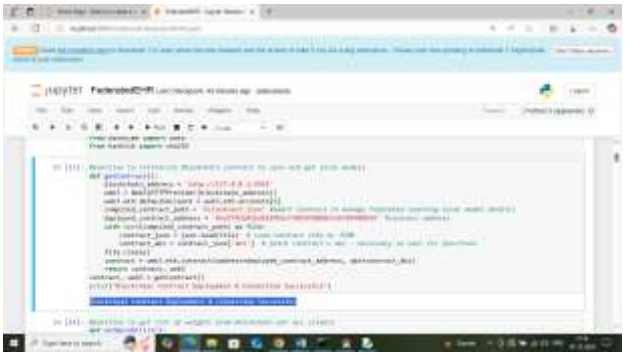
VI. SYSTEM MODEL

System Architecture

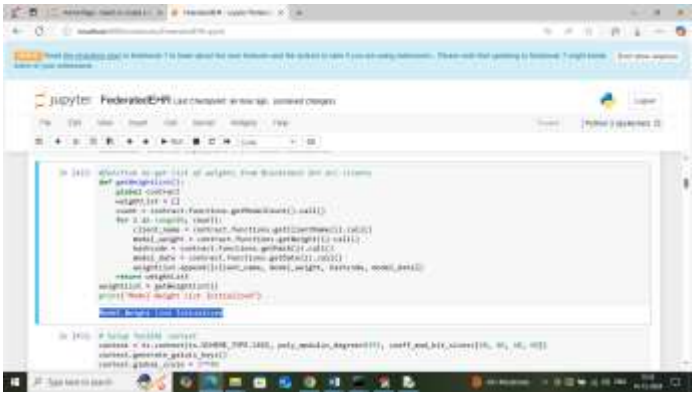


III. RESULTS AND DISCUSSIONS

In above screen importing required python classes and packages



In above screen using smart contract address we are establishing connection with Blockchain to save and get weights



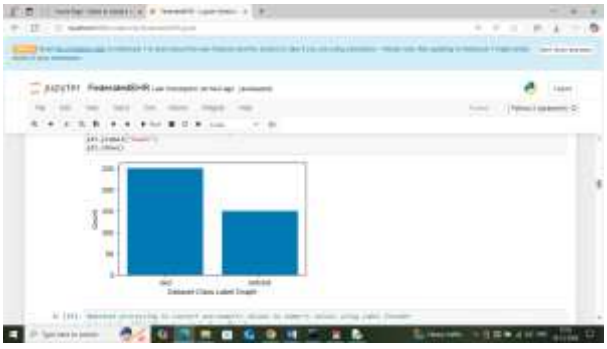
In above screen reading and initializing all available weights list from Blockchain

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

In above screen setting up TEN SEAL context object for Homomorphic encryption using CKKS

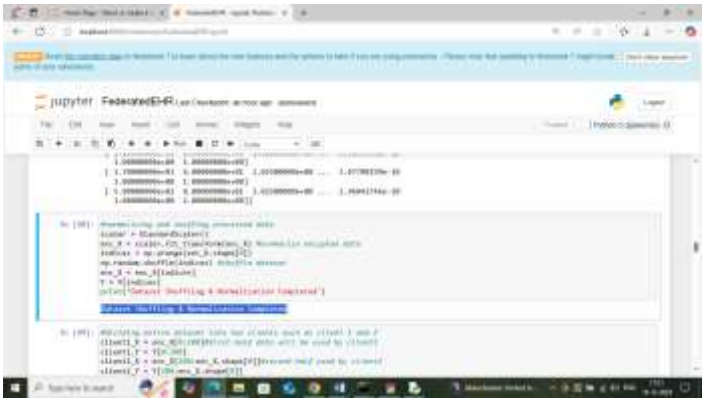
In above screen loading and displaying Chronic Kidney EHR patients dataset



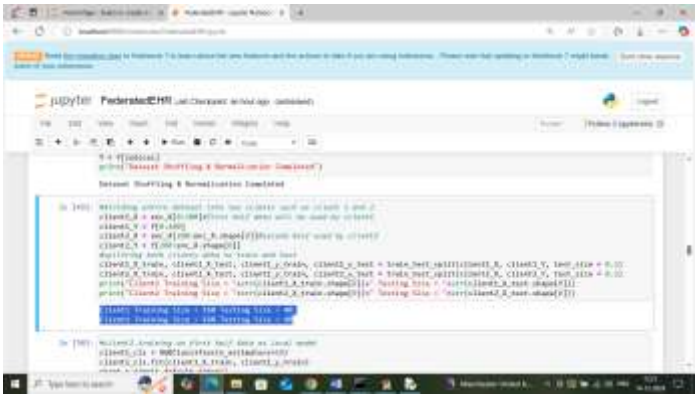
In above screen visualizing graph of CKD and NON-CDK patients where x-axis represents patient type and y-axis represents counts

In above screen applying dataset processing technique such as label encoding to convert all non-numeric data to numeric values and then replacing missing values with mean and then in table can see all values are converted to numeric format

In above screen using CKKS and Homomorphic encrypting all processed dataset and then displaying encrypted data



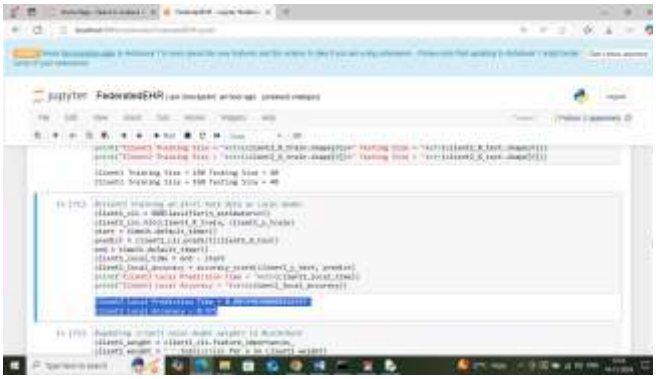
In above screen shuffling and normalizing encrypted data



In above screen dividing entire dataset into two parts where first part used by Client1 and second part used by Client2 and then splitting

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

both clients data into train and test where application using 80% dataset for training and 20% for testing



In above screen training Client1 data using XGBOOST algorithm and then calculating prediction time and accuracy and in above screen Client 1 got 97% accuracy on test data

In above screen extracting weights from Client 1 model and then sending those weights to Blockchain as local model and after storage Blockchain will send output log which is showing in Blue colour text. In above log can see Block no, transaction no, transaction hash and many other details which is a proof of Client 1 weight stored in Blockchain

In above screen training another algorithm on second half data and this model got 92% accuracy and can see computation time also

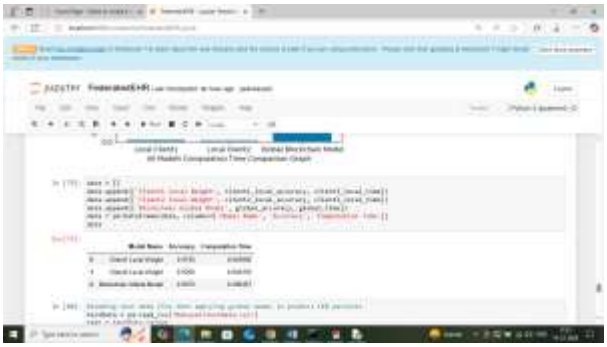
In above screen client 2 model local weight also extracting and sending to Blockchain for storage and then displaying entire log obtained from Blockchain after storage

In above screen generating Global model by merging all local weights from Blockchain and then Global model got 96% accuracy and can see computation time also which is very high

In above screen displaying accuracy comparison graph between all models where client1 local model got high accuracy and Global model got little less accuracy

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

In above screen displaying computation time comparison graph where x-axis represents model type and y-axis represents computation Time and in all models Global model took too much computation time but it will provide high security to data



In above screen displaying accuracy and computation time of all models in tabular format and computation time is in seconds

In above screen reading test data values and then applying CKKS Homomorphic encryption and then applying Global Blockchain model to predict patients disease type and below is the output

In above screen in square bracket can see Test data values and after =➔ arrow symbol can see predicted output as "CKD or non-CKD'.

So above are the output screens of all 3 obejctives.

VIII. CONCLUSION

In this paper, a blockchain-integrated framework for secure federated learning of encrypted Electronic Health Records (EHR) using homomorphic encryption has been proposed to address critical challenges in healthcare data privacy, security, and collaborative analytics. The system effectively combines the strengths of decentralized blockchain technology, privacy-preserving federated learning, and advanced cryptographic techniques to enable secure model training without exposing sensitive patient information.

By ensuring that data remains encrypted throughout the entire lifecycle and eliminating the need for centralized data aggregation, the

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

proposed framework significantly reduces the risk of data breaches, unauthorized access, and single points of failure. The use of blockchain enhances trust, transparency, and traceability among participating healthcare institutions, while smart contracts automate validation and enforce strict access control policies.

The framework demonstrates the ability to achieve high model performance while maintaining strong privacy guarantees, making it suitable for real-world healthcare applications such as disease prediction, clinical decision support, and personalized medicine. Additionally, it promotes secure data sharing and collaboration across organizations without violating regulatory constraints.

Overall, this approach represents a robust and scalable solution for next-generation healthcare systems. It bridges the gap between data utility and data privacy, paving the way for secure, decentralized, and intelligent healthcare analytics. Future enhancements can further optimize computational efficiency and expand its applicability to broader domains within secure distributed machine learning.

IX. FUTURE WORK: Future work for this

While the proposed blockchain-integrated federated learning framework with homomorphic encryption provides strong security and privacy guarantees, several areas can be further explored to enhance its performance, scalability, and real-world applicability. One important direction is the optimization of computational efficiency, as homomorphic encryption introduces significant processing overhead. Future research can focus on lightweight encryption techniques, hybrid cryptographic models, or hardware acceleration (such as GPUs and secure enclaves) to reduce latency and improve system performance.

Another area of improvement is scalability. As the number of participating healthcare institutions increases, the communication and consensus overhead in the blockchain network may grow. Future work can investigate advanced consensus mechanisms, sharding techniques, or layer-2 solutions to ensure efficient operation in large-scale environments.

Enhancing security and robustness is also crucial. Although the current system provides strong protection, it can be extended to defend against more sophisticated attacks such as model poisoning, adversarial attacks, and inference attacks. Incorporating techniques like differential privacy, anomaly detection, and secure multi-party computation can further strengthen the framework.

Interoperability and standardization of healthcare data formats remain key challenges. Future work can focus on integrating

standardized protocols such as HL7 FHIR to enable seamless data exchange between heterogeneous healthcare systems. Additionally, improving compatibility with existing hospital information systems will facilitate easier adoption.

The integration of advanced machine learning models, including deep learning and large-scale AI systems, can also be explored to enhance predictive accuracy. Furthermore, real-time analytics and edge computing capabilities can be incorporated to support time-sensitive healthcare applications such as remote patient monitoring and emergency response systems.

XI. REFERENCES

➤ Satoshi Nakamoto (2008), *Bitcoin: A Peer-to-Peer Electronic Cash System*, introduced the concept of blockchain and decentralized ledgers.

➤ Brendan McMahan et al. (2017), *Communication-Efficient Learning of Deep Networks from Decentralized Data*, proposed federated learning for privacy-preserving distributed model training.

➤ Craig Gentry (2009), *Fully Homomorphic Encryption Using Ideal Lattices*, introduced homomorphic encryption enabling computation on encrypted data.

➤ IEEE (2020), *Blockchain for Secure Healthcare Systems*, discussed blockchain applications in medical data security and sharing.

➤ Elsevier (2021), *Federated Learning in Healthcare: A Survey*, reviewed applications and challenges of FL in medical systems.

➤ IBM (2020), *Blockchain in Healthcare: Opportunities and Challenges*, discussed real-world blockchain healthcare solutions.

➢ J.V.Anil Kumar, Tanguturi Naga Trisha," INTELLIGENT VIDEO CONTENT GENERATION USING DEEP LEARNING", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.044, Page 357 to 364, ISSN:2250-3676, 2025.

➢ J.V. Anil Kumar, Nagella Swarupa Rani," SECURE DATA TRANSMISSION THROUGH HYBRID CRYPTOGRAPHY AND STEGANOGRAPHIC TECHNIQUES", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com,

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

https://doi.org/10.64771/ijesat.2025.046,
Page 373 to 383, ISSN:2250-3676, 2025.