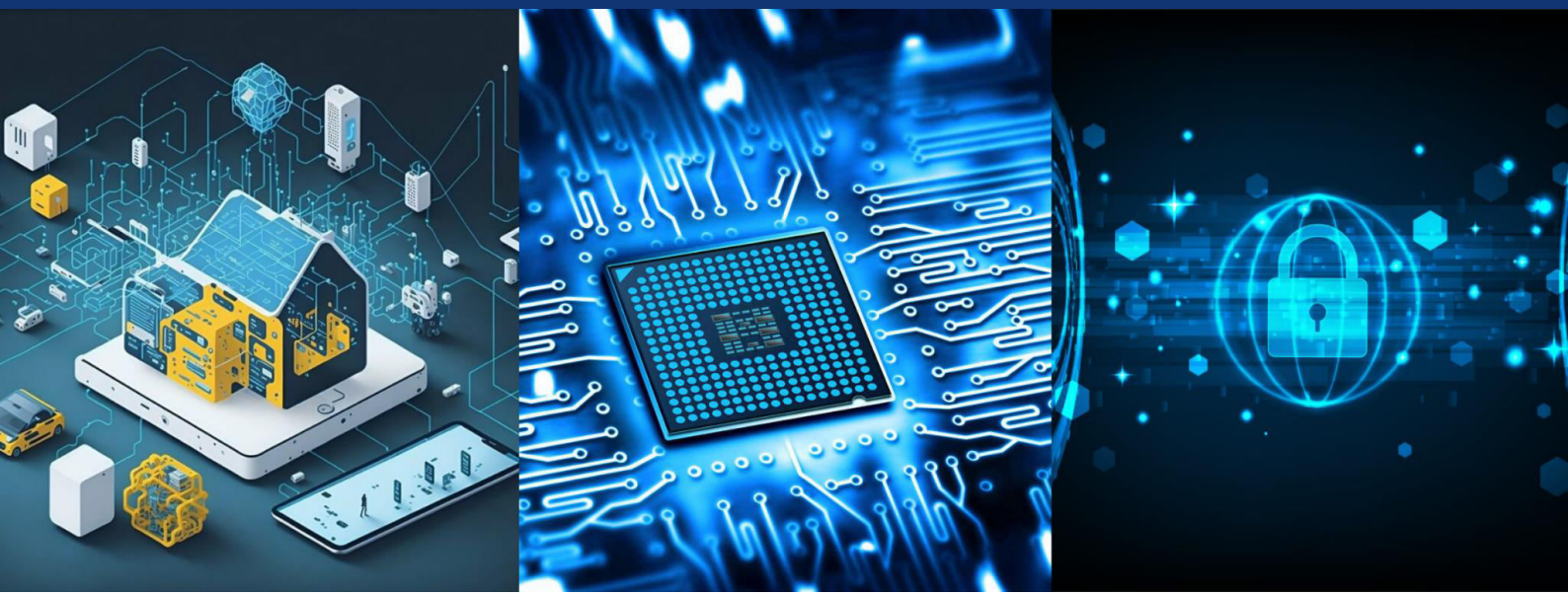
IJIRCCE

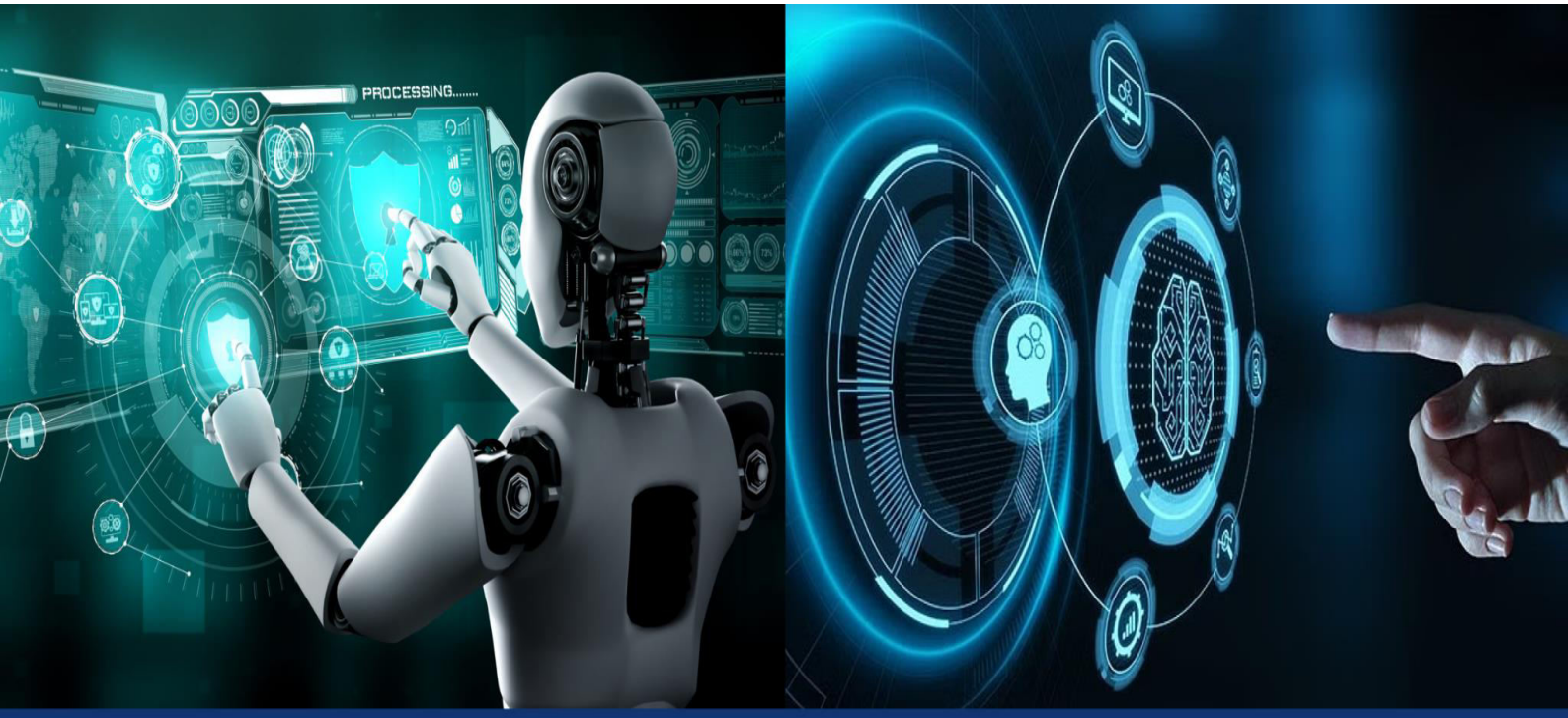
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 7, July 2026

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407058

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption

Dr. R. Sridevi, P. Sanjay Kumar

Professor, Department of Computer Science Engineering, Jawaharlal Nehru Technological University, Hyderabad, Telangana, India

Post-Graduate Student, Department of Computer Science Engineering, Jawaharlal Nehru Technological University, Hyderabad, Telangana, India

ABSTRACT: The rapid growth of cybercrime, ransomware attacks, digital fraud, and large-scale cyber threats has significantly increased the need for secure and collaborative cyber forensic investigations. Traditional machine learning approaches often require organizations to share or centralize sensitive forensic datasets, creating challenges related to privacy, confidentiality, data ownership, and security. To address these limitations, this project proposes a Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption. The proposed framework integrates Federated Learning, Distributed Learning, CKKS-based Homomorphic Encryption, Blockchain Technology, and a Secure Model Exchange Space to enable multiple agencies to collaboratively train machine learning models without exposing their raw forensic data. Federated Learning allows organizations to train models locally and securely aggregate encrypted model updates, while Distributed Learning enables encrypted dataset partitions to be processed collaboratively by helper nodes without revealing the original data. CKKS Homomorphic Encryption protects sensitive information during computation, and blockchain technology provides decentralized trust through secure node authentication, transparent validation, immutable audit trails, and trusted model exchange among participating agencies. The framework is implemented using Python, Flask, Scikit-learn, TenSEAL, Ganache, Solidity, and Web3.py, providing a web-based platform for collaborative project management, encrypted training, blockchain monitoring, secure model sharing, performance evaluation, and cyber forensic prediction. Experimental results demonstrate that the proposed architecture successfully supports secure collaborative learning, encrypted computation, blockchain-based validation, and trusted model sharing while maintaining effective prediction performance. By integrating distributed learning, federated learning, homomorphic encryption, and blockchain into a unified framework, the proposed system provides a scalable, secure, and privacy-preserving solution for next-generation cyber forensic intelligence, enabling organizations to collaboratively strengthen cybersecurity without compromising the privacy, confidentiality, or ownership of sensitive forensic data.

KEYWORDS: Cyber Forensics, Federated Learning, Distributed Learning, Homomorphic Encryption, CKKS, Blockchain, Privacy-Preserving Machine Learning, Secure Model Exchange, Collaborative Learning, Cybersecurity.

I. INTRODUCTION

The rapid advancement of digital technologies has led to a significant increase in cybercrime, ransomware attacks, digital fraud, insider threats, and other sophisticated cyber incidents. As organizations become increasingly interconnected through cloud computing, the Internet of Things (IoT), and distributed information systems, the volume and complexity of cyber forensic data continue to grow. Cyber forensic investigations require the analysis of large amounts of sensitive digital evidence to identify attack patterns, reconstruct incidents, and support legal investigations. However, effective cyber threat detection and forensic intelligence often require collaboration among multiple organizations, making secure data sharing a major challenge.

Traditional machine learning approaches generally rely on centralized data collection, requiring organizations to transfer sensitive forensic datasets to a common server for model training. Such approaches introduce serious concerns regarding data privacy, confidentiality, ownership, regulatory compliance, and the risk of unauthorized access or data leakage.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

These limitations discourage organizations from participating in collaborative cyber forensic investigations and reduce the effectiveness of shared cyber threat intelligence.

Recent advances in Federated Learning, Distributed Learning, Homomorphic Encryption, and Blockchain Technology provide promising solutions for secure collaborative learning. Federated Learning enables multiple organizations to train machine learning models locally while sharing only model updates instead of raw data. Distributed Learning further improves computational efficiency by allowing training tasks to be distributed across multiple helper nodes. Homomorphic Encryption, particularly the CKKS scheme, enables computations to be performed directly on encrypted data, preserving confidentiality throughout the learning process. Blockchain technology establishes decentralized trust by providing secure node authentication, transparent validation, immutable audit trails, and tamper-resistant record management.

This paper proposes a Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption. The proposed framework integrates Federated Learning, Distributed Learning, CKKS-based Homomorphic Encryption, Blockchain Technology, and a Secure Model Exchange Space to enable multiple agencies to collaboratively develop machine learning models without exposing their raw cyber forensic datasets. The framework is implemented using Python, Flask, Scikit-learn, TenSEAL, Ganache, Solidity, and Web3.py, providing secure collaborative training, encrypted computation, blockchain-based validation, trusted model sharing, and cyber forensic prediction. Experimental results demonstrate that the proposed architecture enhances privacy, transparency, security, and collaboration while maintaining effective machine learning performance, making it a scalable and practical solution for next-generation cyber forensic intelligence.

II. RELATED WORK

Several studies have explored privacy-preserving machine learning by integrating blockchain, homomorphic encryption, and distributed learning techniques. R. Yang et al. [1] proposed a blockchain-based federated learning framework that combines homomorphic encryption with a reputation mechanism to improve privacy, security, and trust in distributed learning environments. The framework employs Paillier Homomorphic Encryption, Distributed Gradient Descent, and PBFT consensus to securely aggregate encrypted model updates while mitigating poisoning attacks through reputation-based aggregation. Although the approach enhances collaborative federated learning, it primarily focuses on encrypted model aggregation and does not support encrypted distributed training on partitioned datasets.

P. Pushpa et al. [2] presented a privacy-preserving cloud computing framework based on Partially Homomorphic Encryption (PHE), Somewhat Homomorphic Encryption (SHE), and Fully Homomorphic Encryption (FHE) to protect sensitive data during cloud-based processing. Their work demonstrates secure encrypted computation with advanced key management and end-to-end encryption. However, the proposed framework is limited to secure cloud data processing and does not address collaborative machine learning, distributed model training, or secure model aggregation among multiple organizations. Thanmai Gaddam et al. investigated the use of the CKKS homomorphic encryption scheme and the Concrete ML framework for protecting sensitive healthcare datasets. The study evaluated Logistic Regression, Support Vector Machine, Decision Tree, and XGBoost models using encrypted healthcare data while analyzing encryption overhead and the feasibility of machine learning over encrypted data. However, the work mainly focuses on privacy-preserving inference rather than encrypted model training, collaborative learning, or distributed computation.

Prokash Gogoi and J. Arul Valan [3] proposed a privacy-preserving chronic kidney disease prediction framework that combines Mutual Information feature selection, Logistic Regression, and Paillier Homomorphic Encryption. The model is trained using plaintext healthcare data, while encrypted test data is utilized during prediction to preserve patient privacy. Although the approach achieves accurate predictions while protecting sensitive information, it is limited to secure inference and does not incorporate collaborative learning, blockchain-based validation, encrypted model aggregation, or distributed training. Several other studies have contributed to privacy-preserving distributed learning and blockchain-based systems. W. Fan et al. [4] investigated resource allocation strategies for machine learning-based Industrial Internet of Things (IIoT) applications, while J. Zhang et al. [5] explored federated learning for distributed IIoT intrusion detection. Y. Jiang et al. [6] proposed a blockchain-based data sharing mechanism for IIoT environments, and Q. Li et al. [7] presented a comprehensive survey on federated learning systems, highlighting challenges related to privacy, scalability, and security. Furthermore, M. M. Salim et al. [8] and S. Li et al. [9] demonstrated the effectiveness of homomorphic

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407058

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

encryption for protecting sensitive information in Internet of Medical Things (IoMT) and Industrial Internet of Things (IIoT) applications, respectively.

Despite these advancements, existing approaches generally focus on individual aspects such as federated learning, blockchain, homomorphic encryption, or privacy-preserving inference. Very few studies provide a unified framework that simultaneously supports Federated Learning, Distributed Learning, CKKS-based Homomorphic Encryption, Blockchain-based validation, and secure model exchange for collaborative cyber forensic investigations. To address these limitations, the proposed work introduces a Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption, enabling multiple organizations to collaboratively train machine learning models while preserving data privacy, ensuring decentralized trust, maintaining immutable audit trails, and supporting secure model sharing without exposing sensitive forensic datasets.

III. PROPOSED ALGORITHM

A. *Design Considerations:*

The proposed Privacy-Preserving Distributed Training Architecture is designed based on the following considerations.

- Each participating agency stores its cyber forensic dataset locally without sharing raw data.
- CKKS Homomorphic Encryption is used to encrypt datasets and model updates before transmission.
- Consortium blockchain validates agency registration, training participation, model aggregation, and secure model exchange.
- Federated Learning enables local model training using private datasets.
- Distributed Learning partitions encrypted datasets into multiple chunks that are processed by helper nodes.
- Only authenticated and authorized agencies participate in collaborative learning.
- Blockchain maintains immutable records of registration, validation, aggregation, and model-sharing activities.
- Secure Model Exchange allows only validated models to be shared among authorized agencies.
- The collaborative learning process continues until the global model converges or the predefined number of communication rounds are completed.

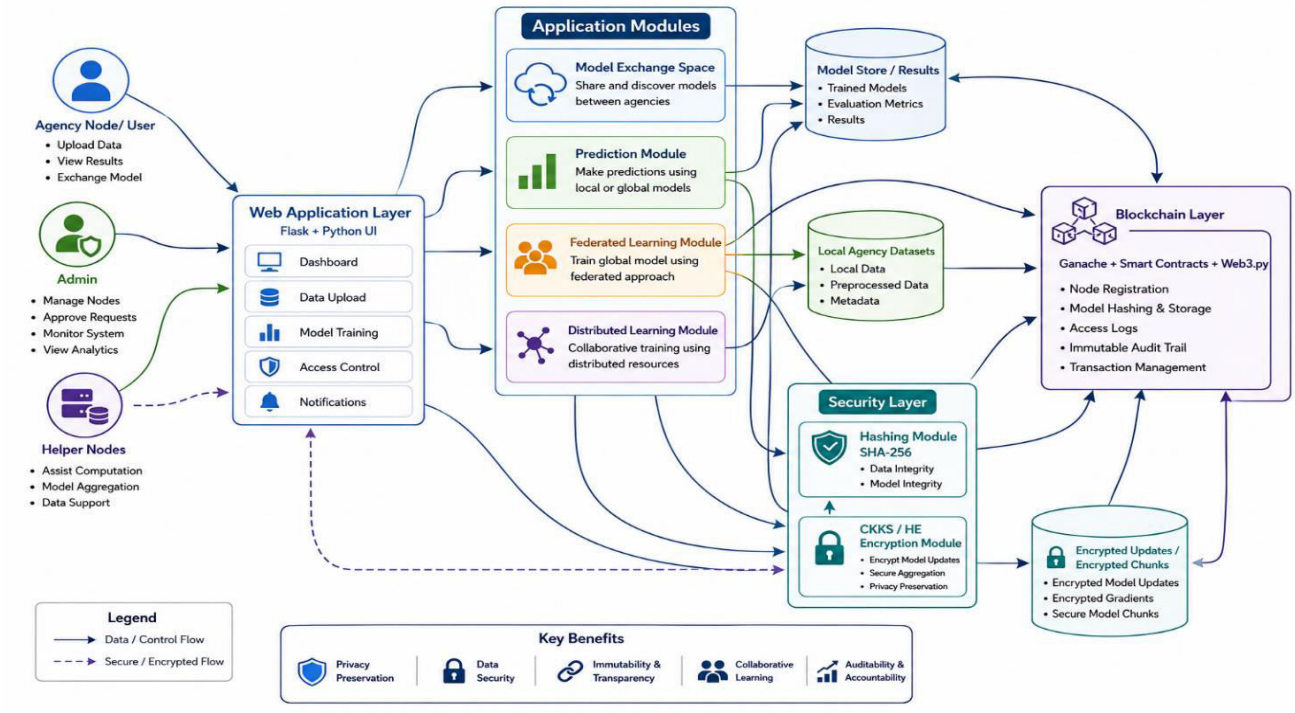


Fig. 1. System Architecture Diagram

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. *Description of the Proposed Algorithm:*

The objective of the proposed algorithm is to enable secure and privacy-preserving collaborative machine learning for cyber forensic investigations without exposing sensitive forensic datasets. The proposed framework integrates Federated Learning, Distributed Learning, CKKS Homomorphic Encryption, Blockchain Technology, and Secure Model Exchange. The algorithm consists of four major steps.

Step 1: Local Model Training and CKKS Encryption

Each participating agency preprocesses its local cyber forensic dataset and independently trains a machine learning model using its private data. The trained model parameters are encrypted using the CKKS Homomorphic Encryption scheme before being transmitted for collaborative learning.

$$M_{enc} = Enc_{CKKS}(M_{local}) \quad ...(1)$$

where,

M_{local} = Locally trained machine learning model parameters.

Enc_{CKKS} = CKKS Homomorphic Encryption function.

M_{enc} = Encrypted model parameters.

Step 2: Blockchain-Based Validation

The encrypted model updates submitted by participating agencies are verified using the consortium blockchain. Smart contracts authenticate participating agencies, validate training requests, and ensure that only authorized encrypted model updates participate in collaborative learning. Every validation event is permanently recorded on the blockchain, providing transparency, integrity, and an immutable audit trail. Only validated encrypted model updates are considered for aggregation, while invalid or unauthorized submissions are rejected.

Step 3: Secure Encrypted Aggregation and Distributed Learning

After successful validation, the encrypted model updates generated by all participating agencies are homomorphically aggregated to construct the global encrypted model.

$$M_{agg} = \Sigma_{(i = 1 \text{ to } N)} M_{enc}(i) \quad ...(2)$$

where,

N = Number of participating agencies.

$M_{enc}(i)$ = Encrypted model update generated by the ith participating agency.

M_{agg} = Aggregated encrypted global model.

For Distributed Learning, the encrypted dataset is partitioned into multiple encrypted chunks and assigned to helper nodes. Each helper node performs computations directly on encrypted data without accessing the original cyber forensic dataset. The encrypted computation results are securely returned and aggregated to generate the encrypted global model.

Step 4: Model Decryption, Evaluation, and Secure Model Exchange

The aggregated encrypted model is decrypted by the project owner using the CKKS secret key to obtain the final global model.

$$M_{global} = Dec_{SK}(M_{agg}) \quad ...(3)$$

where,

Dec_{SK} = CKKS decryption function using the secret key.

M_{agg} = Aggregated encrypted global model.

M_{global} = Final decrypted global machine learning model.

The decrypted global model is evaluated using performance metrics such as Accuracy, Precision, Recall, and F1-Score. If the model satisfies the required performance criteria, it is uploaded to the Secure Model Exchange Space. Blockchain verifies the ownership, integrity, and authenticity of the model before allowing authorized agencies to access it. The collaborative learning process continues until model convergence or the predefined number of communication rounds are completed, thereby providing a secure, scalable, and privacy-preserving framework for collaborative cyber forensic intelligence generation.

IV. PSEUDO CODE

Step 1: Register/Login the participating agencies and create a collaborative learning project.

Step 2: Upload the local cyber forensic datasets, preprocess the data, and partition the datasets for Distributed Learning.

Step 3: Generate the CKKS Homomorphic Encryption keys and encrypt the local datasets/model updates using Eq. (1).

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Step 4: Validate all participating agencies and their encrypted submissions using the consortium blockchain.
Step 5: Check the validation condition for every participating agency.
if (Agency is authenticated AND Model Update is valid)
    Accept encrypted model update
else
    Reject model update
    Record validation failure on blockchain
end
Step 6: Assign encrypted dataset chunks to helper nodes for Distributed Learning and train local models independently using Federated Learning.
Step 7: Aggregate all validated encrypted model updates using homomorphic aggregation according to Eq. (2).
Step 8: Decrypt the aggregated encrypted model using the CKKS secret key according to Eq. (3).
Step 9: Evaluate the global model using Accuracy, Precision, Recall, and F1-Score, and select the best-performing model.
Step 10: Store the validated global model in the Secure Model Exchange repository and record the complete training and model exchange activities on the consortium blockchain.
Step 11: Repeat Steps 3 to 10 until the global model converges or the predefined communication rounds are completed.
Step 12: End the collaborative learning process and display the final prediction results.
Step 8: End.

V. RESULTS

The proposed Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption is implemented using Python, Flask, Scikit-learn, TenSEAL, Ganache, Solidity, and Web3.py. The experimental evaluation is performed using four machine learning models, namely Gaussian Naive Bayes, Logistic Regression, PCA with Ridge Regression, and Ridge Regression. The proposed framework is evaluated under three different environments: Distributed Learning with CKKS, Federated Learning with CKKS, and the Original (Plaintext) implementation. The performance of these approaches is compared based on two important metrics: prediction accuracy and parallel execution time.

The results demonstrate that the proposed privacy-preserving framework maintains prediction accuracy comparable to the original machine learning models while ensuring data confidentiality through CKKS Homomorphic Encryption and blockchain-based validation. Among the evaluated models, Logistic Regression and Ridge Regression achieve the highest prediction accuracy, while PCA with Ridge Regression also provides competitive performance. The results indicate that the encryption process introduces only a negligible impact on prediction accuracy, thereby preserving the effectiveness of collaborative cyber forensic analysis.

The execution time comparison shows that the original implementation requires the least execution time because it performs computations directly on plaintext data without encryption or blockchain validation. The Federated Learning with CKKS framework introduces a small computational overhead due to the encryption and aggregation of model updates. The Distributed Learning with CKKS framework requires comparatively higher execution time because encrypted dataset partitions are processed in parallel by helper nodes before secure aggregation. Although the proposed encrypted frameworks require additional computational time, they provide significant improvements in privacy, confidentiality, secure collaboration, and decentralized trust among participating agencies.

the simulation results demonstrate that the proposed architecture successfully enables secure collaborative model training, encrypted computation, blockchain-based validation, and trusted model exchange while maintaining high prediction accuracy. The framework effectively balances security, privacy, and machine learning performance, making it suitable for collaborative cyber forensic investigations across multiple organizations.

© 2026 IJIRCCE | Volume 14, Issue 7, July 2026 DOI: 10.15680/IJIRCCE.2026.1407058

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

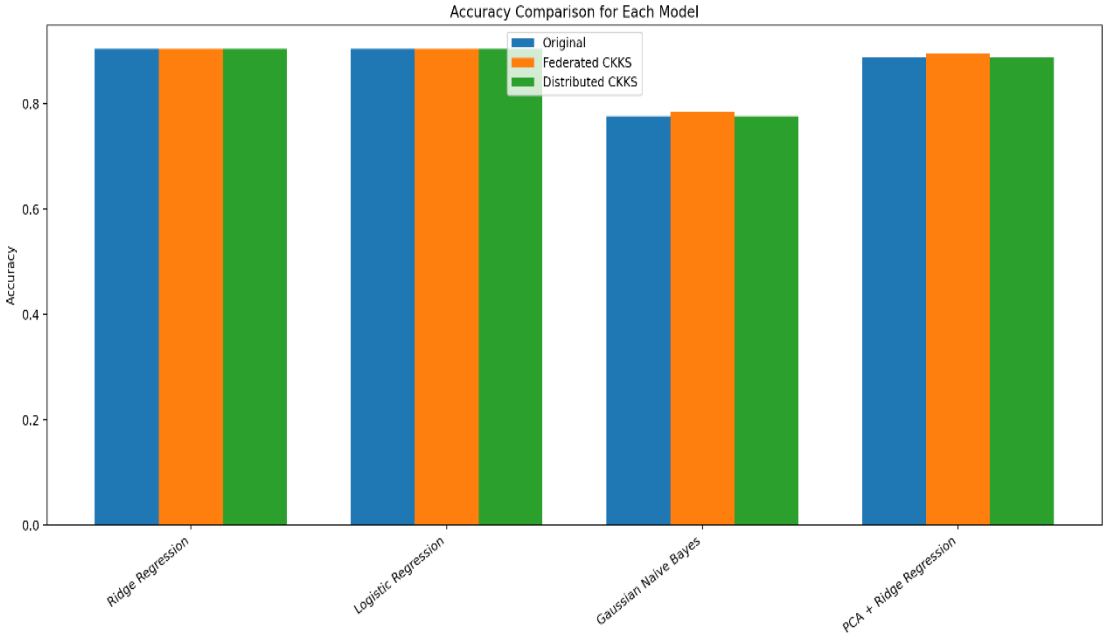


Fig. 1. Accuracy Comparison of Distributed Learning, Federated Learning, and Original Models

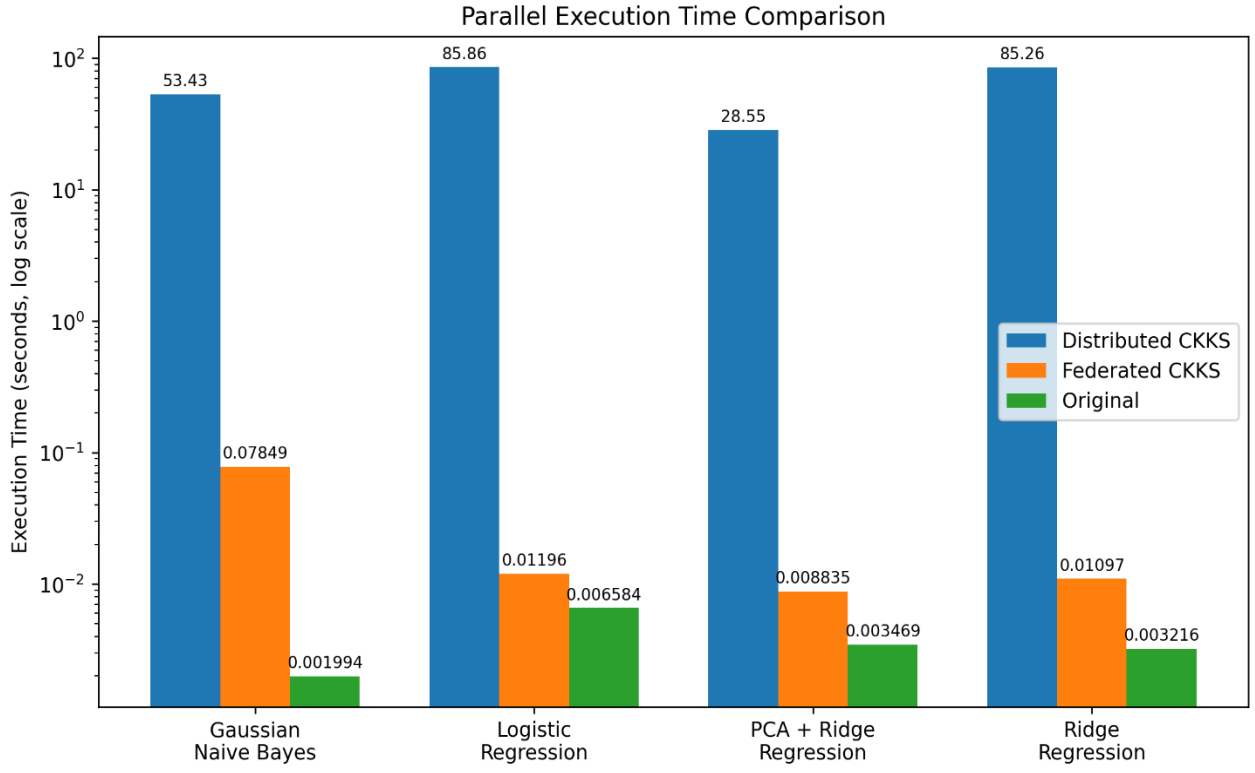


Fig. 2. Parallel Execution Time Comparison of Distributed Learning, Federated Learning, and Original Models

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VI. CONCLUSION AND FUTURE WORK

The proposed Privacy-Preserving Distributed Training Architecture for Cyber Forensics using Blockchain and Homomorphic Encryption provides a secure and scalable framework for collaborative cyber forensic intelligence generation. By integrating Federated Learning, Distributed Learning, CKKS Homomorphic Encryption, Blockchain Technology, and a Secure Model Exchange Space, the framework enables multiple organizations to collaboratively train machine learning models without exposing sensitive cyber forensic datasets. Experimental results demonstrate that the proposed system maintains high prediction accuracy while ensuring data privacy, secure encrypted computation, blockchain-based validation, and trusted model sharing. Although homomorphic encryption and blockchain introduce additional computational overhead, the framework effectively balances privacy, security, transparency, and machine learning performance, making it suitable for collaborative cyber forensic investigations.

Future work will focus on enhancing the framework by integrating advanced homomorphic encryption techniques, secure multi-party computation, and more efficient federated optimization algorithms to reduce communication and computational overhead. The proposed architecture can also be extended to support deep learning models, real-time cyber threat intelligence, cloud-edge deployment, and explainable artificial intelligence (XAI). Furthermore, evaluating the framework on larger real-world cyber forensic datasets with multiple collaborating organizations will improve its scalability, robustness, and practical applicability for next-generation privacy-preserving cybersecurity systems.

REFERENCES

1. R. Yang et al., "Blockchain-Based Federated Learning With Enhanced Privacy and Security Using Homomorphic Encryption and Reputation," IEEE Internet of Things Journal, vol. 11, no. 12, pp. 21674–21688, June 2024, DOI: 10.1109/JIOT.2024.3379395.
2. P. Pushpa et al., "Homomorphic Encryption for Enhanced Cybersecurity Privacy-Preserving Data Processing in Cloud-based Systems," in Proc. 2025 6th International Conference for Emerging Technology (INCET), Belgaum, India, May 23–25, 2025, DOI: 10.1109/INCET64471.2025.11140360.
3. Prokash Gogoi and J. Arul Valan, "Application of Homomorphic Encryption in Machine Learning Based Chronic Kidney Disease Prediction," in Proc. 2024 International Conference (WCONF), July 12–14, 2024, DOI: 10.1109/WCONF61366.2024.10692276.
4. W. Fan, S. Li, J. Liu, Y. Su, F. Wu, and Y. Liu, “Joint task offloading and resource allocation for accuracy-aware machine-learning-based IIoT applications,” IEEE Internet Things J., vol. 10, no. 4, pp. 3305–3321, Feb. 2023.
5. J. Zhang, C. Luo, M. Carpenter, and G. Min, “Federated learning for distributed IIoT intrusion detection using transfer approaches,” IEEE Trans. Ind. Informat., vol. 19, no. 7, pp. 8159–8169, Jul. 2023.
6. Y. Jiang, Y. Zhong, and X. Ge, “IIoT data sharing based on blockchain: A multileader multifollower Stackelberg game approach,” IEEE Internet Things J., vol. 9, no. 6, pp. 4396–4410, Mar. 2022.
7. Q. Li et al., “A survey on federated learning systems: Vision, hype and reality for data privacy and protection,” IEEE Trans. Knowl. Data Eng., vol. 35, no. 4, pp. 3347–3366, Apr. 2023.
8. M. M. Salim, I. Kim, U. Doniyor, C. Lee, and J. H. Park, “Homomorphic Encryption Based Privacy-Preservation for IOMT,” Applied Sciences, vol. 11, no. 18, p. 8757, Sep. 2021, doi: 10.3390/app11188757.
9. S. Li, S. Zhao, G. Min, L. Qi, and G. Liu, “Lightweight Privacy Preserving scheme using homomorphic encryption in industrial internet of things,” IEEE Internet of Things Journal, vol. 9, no. 16, pp. 14542–14550, Aug. 2022, doi: 10.1109/JIOT.2021.3066427.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details