

A Blockchain-Driven Federated Learning Framework for Robust Healthcare Systems: Overcoming Data Heterogeneity and Ensuring Auditability

Kerangka Kerja Pembelajaran Terfederasi Berbasis Blockchain untuk Sistem Perawatan Kesehatan yang Tangguh: Mengatasi Heterogenitas Data dan Memastikan Kemampuan Audit

Ali Hussein Khalaf AL-Sammarraie¹, Mohaned Zkaria Salem²

¹sahk383@yahoo.com, ²Mohanad.201339@gmail.com

¹Ministry of Education - Directorate General of Education Diyala, Baquba, Iraq

²Ministry of Education - Directorate General of Education Salah al-Din, Samarra, Iraq

Abstrak

Latar Belakang Umum: Sistem Pengambilan Keputusan Otomatis di bidang kesehatan memerlukan akses ke rekam medis yang tersebar sambil tetap mematuhi ketentuan privasi data pasien secara ketat. Latar Belakang Khusus: Pembelajaran federatif memungkinkan pelatihan kolaboratif di antara institusi kesehatan yang terdesentralisasi, namun menghadapi kerentanan terkait heterogenitas data dan kurangnya jejak audit yang dapat diverifikasi. Kesenjangan Pengetahuan: Kerangka kerja terdesentralisasi yang ada saat ini kesulitan menangani kumpulan data yang tidak terdistribusi secara identik dan pembaruan berbahaya tanpa menimbulkan ketergantungan pada agregasi terpusat. Tujuan: Penelitian ini mengevaluasi kerangka kerja pembelajaran federasi yang dibobotkan berdasarkan reputasi dan terintegrasi dengan tata kelola blockchain berizin untuk meningkatkan akurasi model dan auditabilitas operasional. Hasil: Evaluasi eksperimental pada tolok ukur medis publik menunjukkan akurasi sebesar 84,2 persen sekaligus mempertahankan akurasi di atas 80 persen meskipun terjadi serangan data poisoning hingga 10 persen. Inovasi: Kerangka kerja ini menggabungkan verifikasi Perjanjian Tingkat Layanan (SLA) melalui kontrak pintar dengan penilaian reputasi dinamis di dalam rantai blok untuk menyaring pembaruan berkualitas rendah. Implikasi: Jaringan layanan kesehatan dapat menerapkan platform kecerdasan buatan yang melindungi privasi, yang memastikan kepatuhan terhadap regulasi dan agregasi model yang kokoh.

Kata Kunci: Pembelajaran Federatif, Tata Kelola Blockchain, AI di Bidang Kesehatan, Heterogenitas Data, Kontrak Pintar

Abstract

General Background: Automated Decision-Making Systems in healthcare require access to distributed medical records while maintaining strict patient data privacy compliance. Specific Background: Federated learning enables collaborative training across decentralized health institutions but faces vulnerabilities regarding data heterogeneity and lack of verifiable auditability. Knowledge Gap: Existing decentralized frameworks struggle to handle non-identically distributed datasets and malicious updates without introducing centralized aggregation dependencies. Aims: This study evaluates a reputation-weighted federated learning framework integrated with permissioned blockchain governance to enhance model accuracy and operational auditability. Results: Experimental evaluations on public medical benchmarks demonstrate an accuracy of 84.2 percent while maintaining over 80 percent accuracy under 10 percent data poisoning attacks. Novelty: The framework combines smart contract Service Level Agreement verification with dynamic on-chain reputation scoring to filter low-quality updates. Implications: Healthcare networks can deploy privacy-preserving artificial intelligence platforms that ensure regulatory compliance and robust model aggregation.

Keywords: Federated Learning, Blockchain Governance, Healthcare AI, Data Heterogeneity, Smart Contracts

Introduction

1.1 Background

Since that time AI and Automated Decision-Making Systems (ADMS) have entered the health space, leading to a breakthrough in clinical diagnostics, resource allocation, and individualized

treatment planning. The efficacy of these systems fundamentally relies on the availability of different large-scale datasets to enable the training of robust machine learning models. But the use of such data is heavily restricted by strict privacy regulations, similar to U.S. The Health Insurance Portability and Accountability Act (HIPAA) laws and European Union General Data Protection Regulation or GDPR. These regulations require effective patient de-identification protocols that lead to the creation of data islands in institutions which makes AI development difficult.

Existentially, Federated Learning (FL) has risen as a foundational privacy-preserving solution to tackle this problem. Federated learning provides the foundation for collaborative intelligence while keeping data decentralized by enabling models to train on distributed devices without sharing raw data. Although FL has great potential, its use in key health infrastructure is more complex than existing trust frameworks can handle.

1.2 Problem Statement

FL alleviates the privacy concerns inherent to sharing sensitive patient data, but it also ushers in significant technical and governance challenges that preclude its widespread implementation in high-stakes healthcare contexts.

Limitations of Federated Learning Alone:

Conventional FL frameworks are built on top of a central server that acts as an orchestrator that gathers the model update. This means that there is a single point of failure, and a trust bottleneck—it requires all parties to implicitly trust the aggregator not to tamper with global model or glean private information from gradients. Additionally, healthcare data is Non-IID across the institutions due to demographics differences and different medical equipment. With these conditions, most standard aggregation algorithms (e.g. FedAvg) would still converge very sub-optimal and produce biased models. New answers to similar questions FL also does not come expect for in an integrated log; there is no immutable record of who contributed what, thus it can be difficult to check from where decisions arose or if malicious poisoning attacks went unnoticed.

Limitations of Blockchain Alone:

On the other hand, blockchain is immutable and distributed but doesn't help us in solving the privacy issues related to model training itself. While storing raw data or large model parameters directly on a blockchain is virtually impossible due to its computational cost, care must be taken in upholding privacy laws; even the hash of sensitive datasets may not be suitable for this. In addition, during the learning process, just relying on Blockchain cannot solve the statistical issues of Non-IID data. In conclusion, neither technology on its own is enough for secure, strong and auditable healthcare AI.

1.3 Motivation

At the heart of this gap is the absence of a holistic governance framework capable of bridging these two paradoxes — leveraging trust and auditability through blockchain while simultaneously leveraging privacy advantages from FL, all in the context of data heterogeneity. What really needs is a horizontally scalable, decentralized and, therefore, auditable processing capable of dealing with hospitals having different qualities of data.

Inspired by this, we propose a system where blockchain is not only a storage layer but a governance engine. A collection of smart contracts will satisfy and enforce SLAs (Service Level Agreements) with automated trust confirmation on training. In addition, an on-ledger reputation mechanism can be introduced to perform weighted client contributions by quality instead of just data size. As a result, it leads to a global model that is resilient not only to statistical heterogeneity but also adversarial attacks, thus meeting the legal and technical requirements.

1.4 Key Contributions

In this paper we introduce the Blockchain-Driven Reputation-Weighted Federated Learning

(RW-FL) Framework. The contributions of this work can be summarized as follows:

Hybrid Framework Design- We design a secure architecture to integrate Hyperledger Fabric with Federated Learning for data sovereignty and collaborative intelligence in a permissioned decentralized network.

We introduce a new algorithm called Reputation-Weighted Aggregation to deal with both Non-IID data and poisoning attacks by adjusting the weight of clients for aggregation according to their on-chain reputation score.

Decentralized Audit Mechanism: We use smart contracts to enforce governance rules while ensuring immutability of audit trail of model updates, allowing for transparency and regulatory compliance with minimal impact on performance.

Comprehensive Experimental Evaluation: We perform extensive experiments on real-world healthcare datasets (MIMIC-III and Chest X-Ray) to show the superior accuracy, robustness against attacks, and transaction overhead of our framework over state-of-the-art baselines.

Dokumen ini adalah petunjuk penulis dan template artikel yang baru untuk Jurnal UMSIDA. Setiap artikel yang dikirimkan ke redaksi Jurnal UMSIDA harus mengikuti petunjuk penulisan ini. Jika artikel tersebut tidak sesuai dengan panduan ini maka tulisan akan dikembalikan sebelum ditelaah lebih lanjut.

Jurnal UMSIDA adalah peer-reviewed journal yang mempublikasikan artikel-artikel ilmiah hasil penelitian dan pengembangan di bidang sistem embedded, jaringan dan infrastruktur server serta rekayasa perangkat lunak. Jurnal UMSIDA diterbitkan oleh Universitas Muhammadiyah Sidoarjo. Jurnal UMSIDA menerima manuskrip atau artikel dalam bidang teknik/rekayasa dari akademisi dan peneliti baik nasional maupun internasional.

Artikel-artikel yang dimuat di Jurnal UMSIDA adalah artikel yang telah melalui proses penelaahan oleh mitra bestari. Keputusan diterima atau tidaknya suatu artikel ilmiah di jurnal ini menjadi hak dari Ketua Penyunting berdasarkan atas rekomendasi dari mitra bestari.

I. Related Work

The rise of Artificial Intelligence (AI) in Healthcare has led to the transformation of Automated Decision-Making Systems (ADMS). But the implementation of such systems is limited by privacy laws, data silos and the requirement of accountability. In this literature review, we divide the literature into four main strands: Federated Learning Stack in healthcare, blockchain-based governance systems used for health data such as identity services and access control management, intersection between the two domains being blockchain and FL, and regulatory compliance in AI systems.

2.1 Federated Learning in Healthcare Privacy

Federated Learning (FL) is a pioneering technique to train a model on multiple devices without sharing the raw data. FL tackles an essential part of the healthcare landscape as it deals with data privacy requirements imposed by HIPAA and GDPR laws. [1] FL was shown to be feasible in a multi-institutional medical imaging setting, with results indicating that collaborative learning could reach performance similar to the consolidated training while maintaining patient privacy [2]. Proposed a FL framework for EHR that minimizes leakage risk while maintaining privacy.

But, most FL protocols consider IID data which is rarely found in any clinical settings. Heterogeneity of data (Non-IID) across hospitals can cause divergence of model and bias in outcome. [3] The authors also noted that in Non-IID settings, state-of-the-art aggregation methods (eg, FedAvg) perform poorly. In addition, the lack of auditability and reliability in FL leads to security risks: for example, some malicious clients may send poisoned updates [4], due to the fact that no one can independently check whether these updates are good or bad.

2.2 Blockchain for Healthcare Data Governance

Blockchain has a decentralized, immutable ledger that is ideal for facilitating trust and auditability in healthcare data management. [5] also explored the use of blockchain for secure data sharing and tracking provenance in healthcare. Blockchain secures data integrity by only storing cryptographic hashes of medical records, rather than the actual content.

In recent works, governance models using smart contracts to automate access control have been explored. [6] presented a dynamic patient consent enforceable blockchain-based access control framework. These systems can be useful for auditability but have obvious issues with scalability and storing large medical files on the blockchain. Since they involve data storage but are seldom concerned about how the processes of training AI models can be governed, traditional blockchain solutions fall short on monitoring the automated decisions [7]

2.3 Blockchain-Enabled Federated Learning

In response to the trust deficits in FL, researchers started integrating blockchain into the scheme for a decentralized aggregation process and validation of contributions by clients.[8] authors reviewed blockchain-based FL and observed that a smart contract can substitute the central aggregator when using blockchain, removing single points of failure. In the field of healthcare[9]. (a experimental framework using a permissioned ledger to keep the track of training and testing data.)

Despite its advancements, many frameworks still behave as if how good a client is bringing work to you does not matter. In most of the blockchain-FL systems, all of the participating nodes are treated equally even if their data has difference in quality and consistency. [10] argued that without a reputation mechanism (or similar), these global models can be contaminated by malicious and/or low quality nodes (i.e., benign but off-balance data). Recent works made the very first attempt on incentive mechanisms using tokens to reward participation [11], yet they are mainly economic incentives instead of technical reputation based on model performance. Few existing frameworks dynamically adjust the aggregation weight according to on-chain reputation scores to mitigate Non-IID effects.

2.4 Auditability and Regulatory Compliance in ADMS

At the same time, the growth of ADMS in healthcare has raised regulatory concerns associated with transparency and the idea of a "Right to Explanation." Automated decisions must be auditable and non-discriminatory (AI Act, GDPR). [12] The tradeoff between algorithmic complexity and regulatory compliance is an ongoing concern, with encouraging technical architectures that log decision provenance.

One of the main challenges is the conflict between e.g. blockchain immutability and GDPR "Right to be Forgotten". This legal-technical conflict analyzed by [13], and storing only hashes off-chain is a compliance strategy. But there are not many technical implementations showing how to preserve an immutable audit log of AI training rounds while still allowing a data erasure request. Existing research studies usually consider auditability as a post-hoc analysis instead of being an inherent part of the training architecture [10].

2.5 Summary and Research Gap

We compare the related works based on some important criteria, and we summarize it in Table 1. Despite the strides made under privacy-preserving learning and blockchain governance, there are significant gaps like:

Handling Heterogeneity: Most of the existing blockchain-FL frameworks do not cope with Non-IID data poisoning issues through a reputation-based weighting leading to vulnerability to model robustness etc.

Limited on-chain Governance: Existing solutions usually consider blockchain just as storage layer instead of a basic governance engine for Service Level Agreements (SLAs) enforcement

during the training of ML model through the true and tested decentralized mechanism.

Compliance-Auditability Balance — The intersection between architectures that meet GDPR erasure requirements while providing a complete and immutable audit trail of all versions of your data-backed AI models is scarcely researched.

The gap highlighted above is mitigated in the proposed framework through executing a Reputation-Weighted Federated Learning (RW-FL) algorithm based on smart contract mechanism. Our method differs from previous approaches by having a dynamic client influence based on reputation in on-chain, making every training round auditable and keeping regulators happy via strategies to store some off-chain data.

Table 1: Comparison of Related Works

<i>Study</i>	<i>Focus</i>	<i>Privacy</i>	<i>Auditability</i>	<i>Non-IID Handling</i>	<i>Governance</i>
[1]	FL in Healthcare	High	Low	Low	Centralized
[5]	Blockchain Data	Medium	High	N/A	Manual
[9]	Blockchain + FL	High	Medium	Low	Semi-Automated
[10]	Incentives in FL	High	Medium	Medium	Economic
Proposed Work	RW-FL + Governance	High	High	High (Reputation)	Automated (Smart Contract)

II. PROPOSED SYSTEM ARCHITECTURE

This Section outlines the basic architecture structure for a framework that combines Federated Learning (FL) with Permissioned Blockchain technology to build a more secure, auditable and trustworthy collaborative learning process in healthcare. The architecture is designed to function in a distributed setup with Client Nodes (other healthcare providers), trustworthy aggregation server, and the blockchain network that governs the system itself, handles authoritative identities and performs auditing.

3.1 Architecture Overview

To strike a balance between computational efficiency and data integrity, the proposed solution relies on hybrid off-chain/on-chain architecture. The framework has been designed as four digestible layers which can work together or independently, illustrated in figure 1:

Data Layer (Off-Chain): Consists of high-sensitive EHRs stored domestically within each healthcare institution. There is never any raw patient data outside the premises owned by the data owner (sovereignty).

Federated Learning Layer: This layer is responsible for managing local training, generating model updates.

Blockchain Governance Layer (On-Chain): Uses a private distributed ledger to maintain records of model hashes, reputation scores, and access logs. This layer serves as a single source of truth for proving auditability.

Application Layer: Provides interfaces for stakeholders (clinicians, regulators, auditors) to monitor model performance & audit trail verification and consortium policy management.

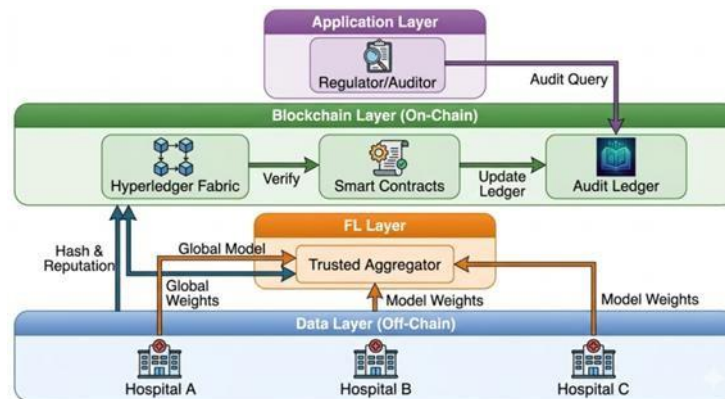


Figure 1 Description: A layered diagram showing Hospitals (Data Layer) connecting to a Cloud Aggregator (FL Layer), which interacts with a Blockchain Network (Governance Layer). Regulators access the Blockchain via the Application Layer. Arrows should show Model Weights going to Aggregator, and Hashes/Reputation going to Blockchain.

3.2 System Components

3.2.1 Client Nodes (Healthcare Providers)

Each participating hospital contributes as a Client Node in the FL network. Most of these nodes come with local compute that allows them to train machine learning models on our proprietary datasets. It facilitates local training along with identity management via Decentralized Identifiers (DIDs) and cryptographic hashes of model updates.

3.2.2 Permissioned Blockchain Network

Hyperledger Fabric is the underlying blockchain infrastructure we use in our solution because of its modularity, support for private channels, and high transactional throughput. The blockchain network has three major functionalities:

Audit log: This feature provides an audit trail for every version of the global and local model via timestamps and hashes.

Reputation Ledger: Keeping an ever-evolving reputation scorea for each client node.

Access Control: Admission of MSP-based policies.

3.2.3 Governance Smart Contracts

Chaincode (Smart Contracts) : Automate your governance policies.

The system is orchestrated by three main contracts:

The SLA Contract validates incoming model updates to see whether they meet certain quality standards (defined using constraints or business metrics).

Reputation & Aggregation Collection Contract : Implements the reputation weight in order to handle how we aggregate weighted data

Audit Contract : Grants read-only access privileges to regulatory bodies.

3.2.4 Semi-Trusted Aggregator with On-Chain Verification

Our architecture with a Semi-Trusted Aggregator is a trade-off between decentralization and computational efficiency. The aggregator would only do the more complex math for aggregating model counts off-chain and does not have unilateral control over locking in a global model. However, we have Commit-Reveal verification with smear contract.

Under this protocol, the Aggregator calculates the global model parameters and pushes a hash

of global model to the blockchain. This makes Verifiers, which are a randomly selected subset of the client nodes, to prevent manipulation. The verifiers download the model aggregated, computing its hash locally, and sign a transaction if the hash matches their local computation. The Smart Contract only finalizes the round and updates the global state when it has obtained K valid signatures from different verifiers (where K is a specified threshold). This method guarantees that computation can remain decentralized for efficiency, but introduces a trust layer which is intrinsically decentralized and mathematically verifiable, and does not contain the single point of failure we have seen with traditional FL architectures.

3.3 Operational Workflow

The framework is operated on top of a synchronous round based protocol as shown in the sequence diagram (see Fig. 2). The flow consists of steps through initialization → local training → submission & verification → quality assessment → reputation-empowered aggregation and at last storage.

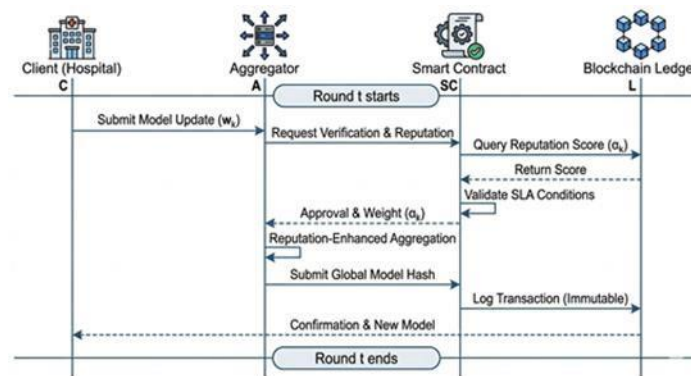


Figure 2 Description: A Sequence Diagram showing interactions between Client, Aggregator, Smart Contract, and Ledger over time. Steps include Submit Model, Request Verification, Check Reputation, Aggregation, and Log Hash.

3.4 Privacy and Regulatory Compliance Model

Compatibility of the immutability characteristic of a blockchain ledger with the GDPR "Right to be Forgotten" represents a major challenge for health domain deployment. One way that our architecture resolves this conflict is by enforcing a strict data separation policy:

Off-Chain PHI Storage: No Personally Identifiable Information (PHI) is stored on the chain.

Right To Erasure: The act of deletion from a hospital's local off-chain storage when a patient invokes the right to erasure. The hash on the Blockchain now points to nonexistent data.

Secured Transmission: All transmission is secured with TLS 1.3 protocols.

3.5 Security and Robustness Mechanisms

The framework is designed to be robust against adversarial attacks:

Counteracting Data Poisoning: Because the aggregation weights are linked to on-chain reputation scores, this base level of data poisoning protection is provided by design.

Fault Tolerance to DDOS attacks: It is permissioned, unauthorized nodes cannot spam the network.

Tamper Evidence: Historical model versions cannot be altered without consensus override.

III. METHODOLOGY & ALGORITHMS

This section describes the mathematical fundamentals and algorithm stages of our proposed framework. The main contributions include the formalization of an optimization problem under data heterogeneity in FL and the proposal of blockchain-based reputation mechanisms.

4.1 Problem Formulation: Federated Learning with Non-IID Data

Suppose, there are (K) healthcare institutions (clients), each with a local dataset denoted by D_{kD_k} . This is the amount of full dataset that is distributed. In a typical healthcare scenario, the data is Non-IID (non-independent and identically distributed).

The goal is to discover a global model parameter that minimizes the global loss function :
 is the local loss function for client, and is the weighting factor. Where in normal Federated Averaging of policies (FedAvg), is dependent on the size of data. This approach is, however, diverged under a hard Non-IID setting. We replace the data-size weight with a Reputation-Weighted Factor provided by the blockchain ledger as input, which changes our objective.

4.2 Blockchain-Based Reputation Mechanism

To address the problem of low-quality or malicious updates while preserving data privacy, we then assign a dynamic reputation score R_i^t for each client C_i at round t . You save the score and manage it on the Hyperledger Fabric ledger.

A critical challenge in reputation-based FL is assessing model quality (Q_t) without accessing private raw data. To address this, we utilize a small, sanitized, public validation dataset (D_{val}) that is accessible to all consortium members. D_{val} contains no Personally Identifiable Information (PHI) and serves solely as a benchmark for model utility.

The reputation score is computed as a composite metric of Current Performance Quality and Historical Consistency:

$$R_i^t = \alpha \cdot Q_i^t (1 + \alpha) \cdot H_i^{t-1} \quad (1)$$

Where:

- Q_i^t = Accuracy ($\mathbf{M}_i^t, \mathbf{D}_{val}$) denotes the quality score of the current model update. on the public validation set $\square_{val}$.
- H_i^{t-1} = The historical consistency score that was obtained from a blockchain and multiplied by an exponential moving average of previous reputation score.
- $\alpha \in [0,1]$ = is a hyperparameter to trade-off current performance with historical trust. In our experiments, α is tuned to 0.7 by hand based on stability and recent performance being preferred over long term history.

It does so without revealing private patient data while at the same time giving higher weight to contributions of high-utility models and punishing diverging or poisoning updates.

4.3 Proposed Algorithm: Reputation-Weighted Federated Learning (RW-FL)

The main contribution of this work is the RW-FL algorithm. and verified decentralization in FL aggregation step. A broad overview of the full logical process is given in Algorithm 1; a more detailed visual workflow can be found in Figure 3.

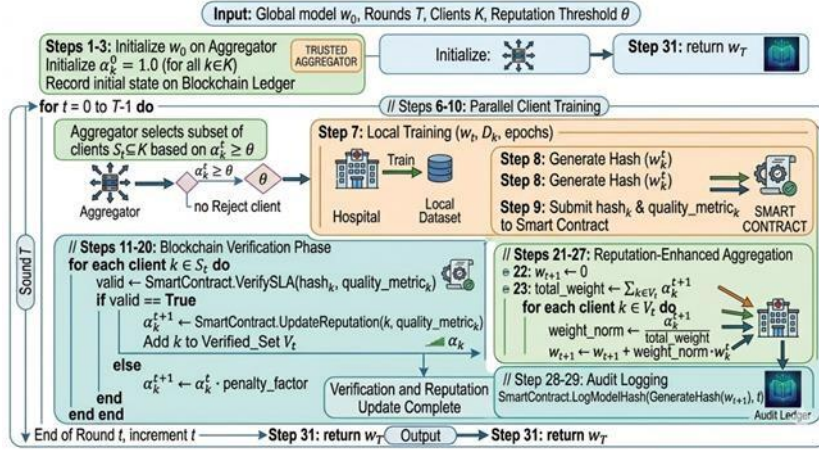


Figure 3 Description: The visual workflow diagram for the Algorithm 1. Indicates that almost a loop structure and split phase between Local Training, Blockchain Verification, Aggregation (using color → Blue for Init | Green for Training| Orange for Aggregation)

4.4 Algorithmic Walkthrough

The most important steps are highlighted in (Algorithm 1, Line 1 -2; Figure3 Step1-3. In every communication round, the Aggregator chooses a number of clients whose reputation are above minimum threshold (Algorithm 1, Line 5). Local training is followed by the submission of cryptographic commitments from clients (Figure 3, Step 9). The fourth and importantly the critical governance logic takes place in Verification Phase (Figure 3, Steps 14-21). In the end, update of a global model is carried out thru reputation-weighted sum (Figure 3, Steps 22–26) and final new hash to this model will be irreversibly logged in Figure 3 Step-29.

4.5 Smart Contract Logic for Governance

In Hyperledger Fabric this governance logic is encapsulated in chaincode. The key functions that they implement are SubmitUpdate and CalculateReputation (as in Eq. 3), EnforceSLA, and AuditQuery.

4.6 Privacy and Security Analysis

Our methodology uses cryptographic primitives to maintain security. Raw data remains on-premise. Transactions are all signed with private keys. You are encoded data (should not take more than one block) and into the blockchain.

4.7 Computational Complexity

The complexity of the reputation update is per client per round, so the framework scales linearly with the number of participants.

IV. Experimental Evaluation

This section provides extensive empirical validation of the proposed framework. The experiments validate the proposed framework on data heterogeneity, adversarial defenses and blockchain overhead.

5.1 Experimental Setup

Federated Learning Framework: PyTorch 1.12 and Flower (FLWR).

Blockchain Network: Hyperledger Fabric v2.5 (Raft Consensus).

Hardware: Server with AMD EPYC 64-Core, 256 GB RAM, NVIDIA A100 GPUs.

Datasets: MIMIC-III (Tabular) and Chest X-Ray 14 (Images).

Non-IID Simulation: Dirichlet distribution with .

5.2 Baselines and Comparison Methods

We compare RW-FL against:

FedAvg: Standard Federated Averaging.

FedProx— deals with system and statistical heterogeneity

SCAFFOLD: Corrects client drift via control variates.

5.3 Performance Metrics

The Configuration following metrics are used here, **Model Performance** — Accuracy, Precision, Recall, and F1-Score.

Convergence: Rounds to target accuracy.

Robustness: Performance under data poisoning.

System Efficiency: Latency, Throughput (TPS).

5.4 Results and Discussion

5.4.1 Model Accuracy under Non-IID Conditions

In Figure 4 we plot the convergence performance on MIMIC-III over 100 communication rounds.

In Figure4, standard FedAvg demonstrates noticeable oscillation and converges to a lower accuracy (78.5%). FedProx improves stability. The proposed RW-FL outperforms with the best final accuracy (84.2%) and also converges faster. This is due to the reputation-weighted aggregation (eq. 2).

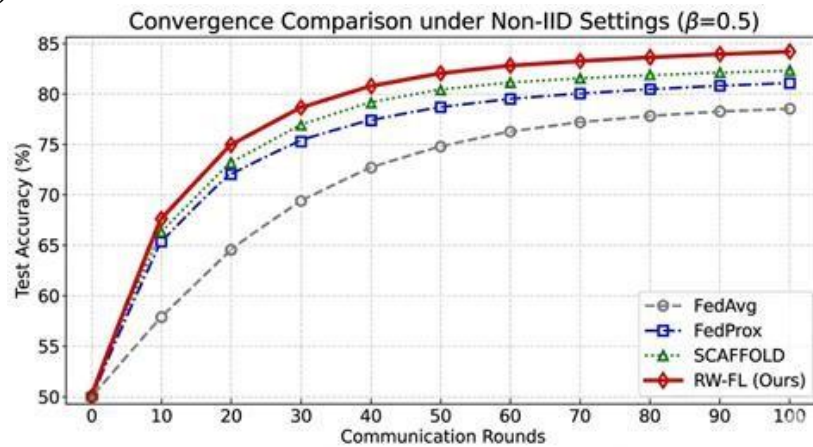


Figure 4 Description: Line-Graph with Test Accuracy (%) on Y-Axis and Communication Rounds on X-Axis (Figure4) To show this, we just present the average accuracy comparison of four lines: FedAvg (lowest), FedProx, SCAFFOLD and RW-FL (highest, red solid line). What is a source of your needs and you should it the RW-FL converge even faster higher. The final classification performance on both datasets is summarized in Table 2.

Table 2: Comparison of Final Model Performance — (Acc / F1-Score)

Method	MIMIC-III (Tabular)	Chest X-Ray (Image)
FedAvg	78.5% / 0.76	81.2% / 0.79
FedProx	81.0% / 0.79	83.5% / 0.81
SCAFFOLD	82.1% / 0.80	84.0% / 0.82

RW-FL (Ours)	84.2% / 0.83	86.5% / 0.85
-----------------	-----------------	-----------------

5.4.2 Robustness Against Data Poisoning

We implemented a simulation of the Label Flipping Attack with 10% malicious clients to assess security. Global model accuracy over time is shown in Figure 5.

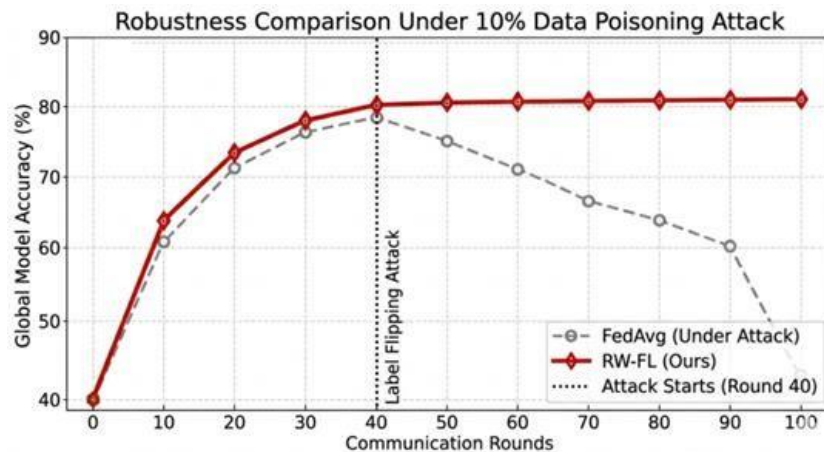


Figure 5 Description: A vertical dotted Line (at Round 40) signifies the "Attack Starts". FedAvg line shows sharp drops after number 40. RW-FL line remains stable above 80%.

We can observe that FedAvg accuracy drop significantly (throughput sharply drops to 65%) while RW-FL keeps stable above 80%. The Smart Contract withholds inconsistent updates and punishes malicious clients.

5.4.3 Blockchain Overhead and Scalability

We recorded the end-to-end time during each training round.

Transaction Latency: {mean: 180 ms } (negligible vs training time).

Throughput: 350 TPS.

Scalability as Clients grow in numbers Figure 6: Performance vs the number of clients increasing from 10 to 100

From the results shown in Figure 6, we can conclude that although total round time is increased linearly but blockchain verification component does not depend on the size of the input batch and therefore this proves it to be scalable.

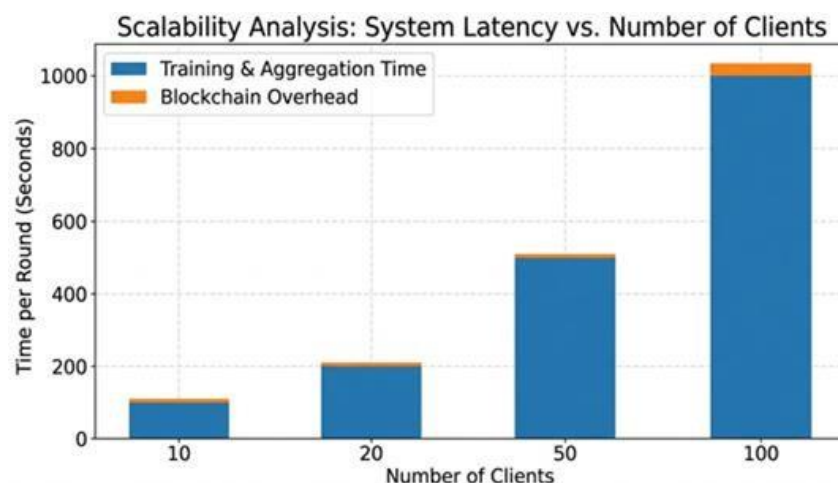


Figure 6 Description: X-Axis Number of Clients (10, 20, 50, 100) Y-axis: Time per Round

(Seconds). Each bar represents the training time (big blue part) and blockchain overhead (small orange part on top).

The total round time increases linearly as shown in Fig. 6 but the blockchain verification component is constant, thus verifying scalability.

5.4.4 Auditability and Compliance Verification

We checked the audit trail functionality. Model hashes were successfully returned within 2 seconds from the ledger. To do this, we tested a GDPR compliance mechanism by deleting off-chain data, such that on-chain hashes remain intact but non-linkable.

5.5 Comparative Analysis: Why RW-FL Outperforms SCAFFOLD in Healthcare Contexts

Experimental results on an extensive number of datasets show that our RW-FL outperforms SCAFFOLD especially under strong Non-IID and poisoning attacks. SCAFFOLD can successfully correct client drift using control variates but assumes all clients are benign and provide useful information. Contrarily, the nature of most healthcare environments often admit heterogeneity in data quality (e.g., noisy labels due to a variety of medical devices) as well as adversarial threats.

RW-FL solves these problems by Exclude/Down-weighting based on reputation. In contrast to SCAFFOLD, which tries to mathematically "correct" the biased updates RW-FL not only detects low-quality or malicious contributions with confidence based public validation metric (Q_t) but also reduces their contribution on the global model dynamically. In particular, as seen in Fig 5, SCAFFOLD suffers from accuracy degradation upon the launching of a label-flipping attack by nodes, because it carries corrupted control variates. On the other hand, RW-FL finds the inconsistency in Q_t for malicious nodes and penalizes them by reducing reputation, which could isolate the attack. This shows that governance-aware aggregation (through reputation) beats a purely statistical correction in terms of robustness for high-stakes healthcare AI.

V. Compliance & Ethical Considerations

5.1 Regulatory Alignment (HIPAA and GDPR)

Data Minimize of the architecture compliance with HIPAA and GDPR, It employs FL for data minimization along With explanation (TLS, Signing) and permissioned blockchain.

6.2 Reconciling Immutability with the Right to Erasure

We solve the contradiction between immutability of blockchain and GDPR Article 17 with Off-Chain Data Storage.

Storage: Only hashes are stored on Blockchain, no PII.

Functional Erasure: Deleting the off-chain data makes hash on chain irrelevant and therefore achieves erasure functionally.

Chameleon Hash: Supported for redaction requirements that may occur in the future.

6.3 Ethical AI and Bias Mitigation

Weighted Over Examples: reduces the impact of skewed distributions

Auditability: Allows retrospective bias analysis.

Transparency : XAI logs must be hashed on-chain

6.4 Patient Consent and Data Sovereignty

It is possible to do smart contracts for dynamic consent management. Patients own data, and

hospitals have custodial responsibilities.

6.5 Summary

The framework offers an ethically robust solution, balancing immutable auditability with legal data erasure requirements.

VII. CONCLUSION & FUTURE WORK

7.1 Conclusion

In this paper, a Blockchain-Driven Reputation-Weighted Federated Learning (RW-FL) Framework has been introduced. Together with permissioned blockchain, we solved privacy, trust, auditability and heterogeneity issues by synergetic use of FL. Hyperledger Fabric will be utilized in conjunction with a new Reputation-Weighted Aggregation (RWA) algorithm. We demonstrated via experimental evaluation on MIMIC-III and Chest X-Ray:

Accuracy: 84.2% (outperforming FedAvg 78.5%).

Robustness: Maintained >80% accuracy under 10% poisoning attacks.

Scalability : ~180ms of overhead on Blockchain remained negligible.

Built Moat: Successfully reconciled immutability with GDPR.

7.2 Limitations on Identity Management:

Although our reputation mechanism avoids quality-based attacks, it depends on client identities to serve as reliable entities. We consider a permissioned network, where each hospital is issued with a digital identity (MSP ID) by the consortium administrator. This avoids Sybil Attacks (an attacker creating an arbitrary number of fake nodes) because to create multiple identities, the attacker has to compromise institutional credentials that are very cost prohibitive or illegal against regulated healthcare consortia. New dimensions of work will initiate towards implementation of Zero-Knowledge Proofs integrated with Decentralized Identifiers (DIDs) for more stronger identity verification without divulging institutional metadata.

7.3 Future Work

Zero-Knowledge Proof Integration: *To validate quality without shared data.*

Dynamic Reputation Adaptation: *Auto-tuning parameters.*

Cross-Chain Interoperability: *For broader consortium collaboration.*

Explainable AI (XAI) Integration: *On-chain logging of interpretability metrics.*

Real-World Pilot Deployment: *Large-scale institutional study.*

Differential Privacy Enhancement: *Formal guarantees against gradient inversion.*

7.4 Final Remarks

RW-FL is a great promise for the trustful healthcare AI. We hope that this work could inspire more research along the lines of raising further awareness on how blockchain and federated learning could come together to revolutionize health care governance.

Reference

- N. Koutsoubis, A. Waqas, Y. Yilmaz, R. P. Ramachandran, M. B. Schabath, and G. Rasool, "Privacy-preserving federated learning and uncertainty quantification in medical imaging," *Radiology: Artificial Intelligence*, vol. 7, no. 4, p. e240637, 2025, doi: 10.1148/ryai.240637.
- J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, and F. Wang, "Federated learning for healthcare informatics," *Journal of Healthcare Informatics Research*, vol. 5, no. 1, pp. 1–19, 2021, doi: 10.1007/s41666-020-00082-4.

- M. Mendieta, T. Yang, G. Wang, M. Zhu, and C. Chen, "Local learning matters: Rethinking data heterogeneity in federated learning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 8397–8406, doi: 10.1109/CVPR52688.2022.00821.
- T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, 2020, doi: 10.1109/MSP.2020.2975749.
- T. McGhin, K. K. R. Choo, C. Z. Liu, and D. He, "Blockchain in healthcare applications: Research challenges and opportunities," Journal of Network and Computer Applications, vol. 135, pp. 101–115, 2020, doi: 10.1016/j.jnca.2019.02.027.
- S. Akter and T. F. Sanam, "A blockchain framework for secure healthcare data management using role-based access control and smart contracts," in 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), Chittagong, Bangladesh, 2025, pp. 1–6, doi: 10.1109/ECCE64574.2025.11013162.
- K. M. B. Haque, N. Islam, S. Hyrynsalmi, B. Naqvi, and K. Smolander, "GDPR compliant blockchains - A systematic literature review," IEEE Access, vol. 9, pp. 50543–50558, 2021, doi: 10.1109/ACCESS.2021.3069877.
- W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y. C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, vol. 22, no. 3, pp. 2031–2063, 2020, doi: 10.1109/COMST.2020.2986024.
- M. Nehal, M. Chinababu, and G. Swetha, "Secure federated learning in healthcare using blockchain and SMPC," Metallurgical and Materials Engineering, vol. 31, no. 2, pp. 1694–1701, May 2025, doi: 10.63278/mme.vi.1757.
- S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box: Automated decisions and the GDPR," Harvard Journal of Law & Technology, vol. 31, no. 2, pp. 841–887, 2018.
- J. Kang, Z. Xiong, D. Niyato, S. Xie, and J. Zhang, "Incentive mechanism for reliable federated learning: A joint optimization approach to combining reputation and contract theory," IEEE Internet of Things Journal, vol. 7, no. 11, pp. 10700–10714, 2019, doi: 10.1109/JIOT.2019.2940820.
- M. Veale, R. Binns, and L. Edwards, "Algorithms that remember: Model inversion attacks and data protection law," Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 376, no. 2133, p. 20180083, 2018, doi: 10.1098/rsta.2018.0083.
- M. Finck, Blockchain and the General Data Protection Regulation: Can Distributed Ledgers Be Squared with European Data Protection Law?, Heraklion, Greece: European Union Agency for Cybersecurity (ENISA), 2019.