THE ASIAN BULLETIN OF BIG DATA MANAGMENT
Vol. 5. Issue 3 (2025)
https://doi.org/10.62019/abbdm.v5i3.373

Blockchain-Enabled Federated Learning for Privacy-Preserving AI Applications

Muhammad Talha Tahir Bajwa*, Muhammad Zeeshan Shafi, Muhammad Atta Ur Rehman, Asad Ali, Faizan khawar, Muhammad Awais

Chronicle

Article history
Received: June 24, 2025
Received in the revised format: July 28, 2025
Accepted: Aug 18, 2025
Available online: Aug 29, 2025

Muhammad Talha Tahir Bajwa, Muhammad Atta Ur Rehman, Asad Ali, Faizan khawar, Muhammad Awais are currently affiliated with the University of Agriculture Faisalabad Department of Computer Science, Pakistan.
Email: talhabajwa6p@gmail.com
Email: attaurrehmanbrw789@gmail.com
Email: imasad34@gmail.com
Email: faizankhawar327@gmail.com
Email: qarni.brothers786@gmail.com

Muhammad Zeeshan Shafi are currently affiliated with the The Islamia University of Bahawalpur, Department of Computer Science, Pakistan.
Email: m.zeeshan.shafie@gmail.com

Corresponding Author*

Abstract

The seamless adoption of the artificial intelligence (AI) in some sensitive areas has heightened discomforts on the privacy of data, security and confidence. Another alternative is Federated Learning (FL), an alternative dedicated to centralized model training and data localization prevention, therefore, eliminating exposure risks. Despite these advantages, FL is also marked by such vulnerabilities as model poisoning, unreliability aggregation, and centralized-coordinator dependency. This paper outlines the current difficulties, and presents a federated learning system with a blockchain-based solution to the current issues - via immutability, decentralization, and transparent use of blockchains. Smart contracts will be used to govern the secure verification of participants, fair incentive systems and the impossibility to tamper with updates to models. Such an integration will allow trustless cooperation between untrusted parties and protect sensitive data. The suggested framework is evaluated in a variety of application permutations of AI, such as healthcare, finance, and IoT environments. Experimental results show that by means of blockchain-enabled FL, the privacy-preserving property is represented, robustness against malicious attacks is boosted and scaleable and trust-sensitive AI solutions are developed. The converge opens the path to sustainable, privacy-preserving and secure AI applications in the real world.

Keywords: Blockchain, Federated Learning, Privacy-Preserving AI, Decentralized Systems, Smart Contracts, Security, Internet of Things (IoT).

INTRODUCTION

An increasing number of AI applications are deployed today in such important domains as healthcare, finance, education, and Internet of Things (IoT) [1], [2] due to the convergence of computational methods efficiency and massive amounts of digital data. Increasingly, AI-based solutions are used in the decision-making process, predictive analytics, and semi- and fully autonomous process execution. Yet the use of such systems is limited by the necessity of large, high-quality datasets [3]. There are dangers in traditional centralized machine learning structures. The centralization of sensitive data in one location on a server increased chances of getting damaged by viruses, risked fraud by unauthorized users and a threatened breach of regulations [4]. As an example, in the healthcare sector, institutions may be reluctant to share patient data between organizations to comply with confidentiality requirements of various legal acts such as GDPR (2016) and HIPAA (1996) [5]. Similarly, financial bodies challenge the norm of sharing transactional information to detect frauds without loosing client confidence [7]. The above problems underline the urgency of the need to create decentralized solutions that do not circumvent privacy but also allow collaborative intelligence [6]. FL is quickly becoming an attractive measure to the

problems mentioned. FL minimizes the risks of data centralization by its training models locally and only transmitting the updated models instead of data transfer [8]. It enables the joint action of many organizations or devices to jointly advance global models without interfering with the data autonomy of any organization. In addition to these gains, FL has weaknesses that include overreliance on central aggregator, model poisoning, and unverifiably auditable process [9], [10]. These deficiencies restrict its use in trust-sensitive applications in the real world.

The issue of FL can be addressed via the combination of blockchain technology that has recently been proposed. Due to its immutability, lack of centralization, and transparency, it is presented as the promising way to enhance trust in distributed learning system [11]. By applying blockchain to FL, the participants will be able to cooperate with each other without having to trust an authority. Smart contracts make this integration even more effective as they automate functions like verification of participants, distributions of incentives and secure logging of the updates of a model [12]. Combined, blockchain-enabled federated learning (BFL) provides a potentially superior model of developing privacy-preserving, secure and trustworthy AI applications.

Artificial Intelligence (AI)

Artificial Intelligence (AI) can be described as the imitation of human intelligence by the machines which have the capability to do work like learning, reasoning, problem solving, sensing and understanding of natural language. According to [19], I can define IA as a study of those agents which perceive the environment and act upon it to increase the likelihood of accomplishing objectives. Its use has spread to a dramatically innovative entity in various fields: healthcare, finance, transportation, cybersecurity, and the Internet of Things (IoT). Its uses include medical image processing and disease forecasting, as well as fraud detection, recommender systems and autonomous driving. The popularity of AI is mainly induced by the increased amount of data, growth of computational capabilities, as well as specialized algorithms.

Machine Learning (ML) and Deep Learning (DL)

Machine Learning (ML) is a branch of AI dealing with the same approach to creating algorithms that could learn without being directly instructed. The models themselves act similarly to ML models in that they enhance their performance with each exposure to new data [20].

Types of ML:

- **Supervised Learning:** Model learns on labeled collections of data in order to make predictions.
- **Unsupervised Learning:** Models use un-labeled information and identify the concealed schemes.
- **Reinforcement Learning:** Modeling is based on trial and error interactions with the environment.

DL is a subsequent sub-discipline of ML that employs artificial neural networks (ANNs) that have several layers to represent complex non-linear relationships. In the fields of image recognition, natural language processing and speech recognition, deep learning algorithms (convetional neural networks CNNs and recurrent neural networks RNNs) are being utilized widely [21].

Limitations of Centralized Machine Learning

Conventional ML/DL models rely on the centralization of data necessitating the need to transmit raw information on distributed sources to a central server to train it. Although a successful strategy in terms of the model accuracy, this has major limitations:

- **Data Privacy Issues:** It may be sensitive data (medical records, financial transactions) that are exposed in the transfer process.
- **Security Risks:** Servers centralize and become prime targets to computer attacks and data breaches.
- **Regulatory Compliance**: Policies such as the GDPR and HIPAA limit the collection, storage and transfer of personal data to specific ways.
- **Scalability Bottlenecks:** During the transfer of the massive amounts of raw data across networks, there would be a lot of bandwidth consumptions and delays.

The restrictions spurred the innovation of Federated Learning (FL) as an alternative that was more protective of privacy.

Federated Learning (FL)

Federated Learning is a distributed learning process that collaboratively trains that global model by retaining the raw data with actors (e.g. cell phone, hospital, bank) who retrain the global model locally on their systems. Instead, everyone builds the model locally and only uploads model changes (gradients or parameters) to a remote server, which adds it all together to generate a model update.

The typical workflow of FL is the following:

- A global model is initialized on a central server
- Clients download the model and perform training using their proprietary data on their own end.
- Clients also share encrypted or masked updates of their models with the server.
- The server then averages these updates (often, using federated averaging, FedAvg).
- The new model is resold to clients; the iteration is repeated.

This de-centralized model minimizes the risks of data exposure and helps to adhere to the data privacy rules.

Types of Federated Learning

Depending on how the data is distributed across participants, there are various types of L can be described as such:

Horizontal Federated Learning (HFL)

- The clients are handled together in the feature space but different samples are attributed to them.
- Example In hospitals in other regions, patients also have the same attributes (age, blood pressure, test results).

Vertical Federated Learning (VFL)

- The sample IDs are concordant across clients; however, the feature spaces are different

- An example is that a hospital and a bank have information on the same person yet the characteristics are varying (health-related versus financial).

Federated Transfer Learning (FTL)

- The samples and the feature spaces are different among participants.
- Transformation learning is used to fill in the gaps that exist across heterogeneous data sets.

Each of them deals with a different real-world scenario of data distributions, which makes FL flexible with respect to collaborative AI.

Challenges in Federated Learning

Although FL has very favorable privacy aspects, it continues to be burdened by some major challenges:

- **Central Aggregation Risks:** Involves the use of a common server that causes the single point of failure.
- **Vulnerability to Poisoning Attacks:** Bad clients would post corrupted updates to model.
- **Unreliable Participants:** Free-riders might spend little or nothing useful on the system, yet make use of it.
- **Absence of Transparency:** There is lack of opportunity to identify dishonest conduct without the ability to prove its true nature through auditing [22].

Such restrictions restrain the use of FL in very sensitive areas.

Blockchain Technology

Blockchain is a decentralized and distributed type of ledger whereby the transactions are recorded as blocks which are chained cryptographically. A block is composed of a list of transactions, which are confirmed by using consensus (e.g. mechanism Proof of Work, Proof of Stake) [23].

Key Features

- **Immutability:** Information can never be altered.
- **Decentralization:** There is no chief authority over the network.
- **Transparency:** The transactions can be checked by all the participants.
- **Security:** Cryptographic techniques prevent tampering and unauthorized access.

Types of Blockchain

There are four types of blockchains:

- **Public Blockchain:** Available to selected participants (e.g., Bitcoin, Ethereum).
- **Private Blockchain:** Accessible to a small group (e.g., Hyperledger Fabric).
- **Consortium Blockchain:** It is also managed by one organization, but also cases where more than one organization makes decisions.
- **Hybrid Blockchain:** Hybrid Blockchains have the properties of both a public and a private blockchain.

Whether to use blockchain or not is dependent on the type of application, scalability demands and the assumptions about trust.

Privacy-Preserving AI

Privacy-preserving AI are methods that enable training and predictions of machine learning models so that sensitive data are not revealed. Methods include:

- **Differential Privacy (DP):** Adding noise of data updates or model updates to make minimized exposure of information leakage potential [24].
- **Homomorphic Encryption (HE):** Homomorphic computation in encrypted area without decryption [25].
- **Secure Multi-Party Computation (SMPC):** Computation over a publicly shared dataset that is divided among several parties [26].

These approaches improve on privacy but suffer computation overheads and scalability. Blockchain-based FL has the capacity of offering an alternative as it ensures transparency, decentralized trust and secure handling of data.

Blockchain-Enabled Federated Learning (BFL)

The use of blockchain in FL results in Blockchain-enabled federated learning (BFL), which mitigates the weaknesses of conventional approaches of FL. In BFL:

- Blockchain eliminates the single point of failure as blockchain then replaces the central aggregator with a decentralized ledge.
- The use of smart contracts will automatize such processes as verifying the participants, distribution of the incentive, and auditing of the model updates.
- Transparency is the guarantee, that an evil activity (such as poisoning attacks) can be detected, and then recorded for eternity.

Such synergy will make it possible to have trustless cooperation among untrusted parties, improving privacy and security of AI applications.

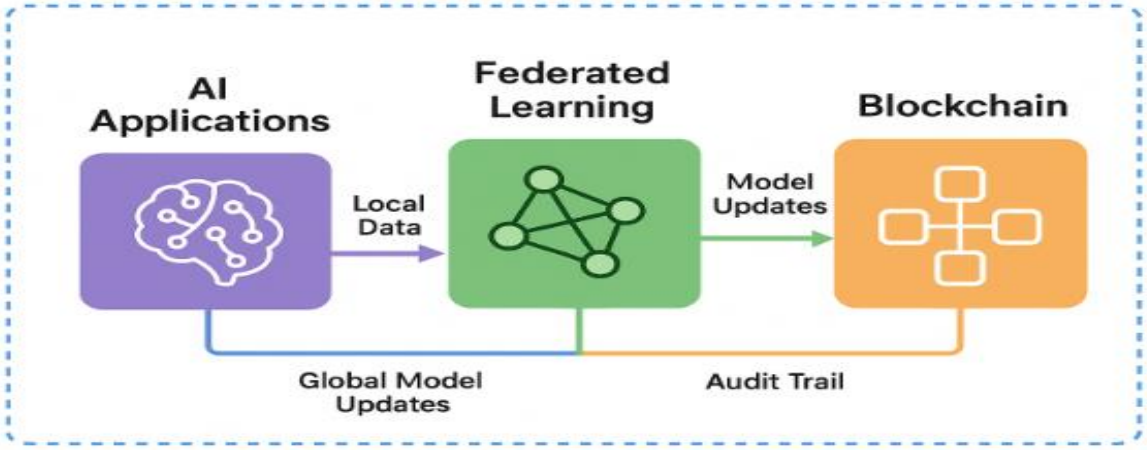

Figure 1.
Conceptual workflow of Blockchain-Enabled Federated Learning (BFL)

Federated learning nevertheless has its flaws in that it is still prone to central aggregation, malicious update problems, and has no transparent auditing. The existing solutions fail to ensure unparticipant trust to a full extent or the integrity of the training process verification. The need therefore exists to have a decentralized, tamper-free and scalable framework that can further promote privacy conservation without impairing the strength or the trust of data collaboration in AI systems.

Research Contributions

The present paper has the following contributions:

- **Framework Design -** The proposed blockchain-enabled federated learning (BFL) framework provides a combination of blockchain components that offer decentralization, transparency and smart contracts automation, with the privacy benefits of FL.
- **Security and Trust Mechanisms -** Presents mechanisms to authenticate securely model participants, prove the integrity of aggregations and in an open manner audit the model upgrades.
- **Incentive Management -** Creates strategic contract-based incentive programs that are used to reward the pursuit of honest action and penalize malicious behavior.
- **Domain Applicability -** Tests the transformation capability of BFL upon assessment of major areas of I within the context of health, finance and IoT environments.
- **Performance Analysis -** Analyzes the trade-offs between blockchain integration and FL efficiency, with an emphasis on scalability, latency, robustness of applications in the real world.

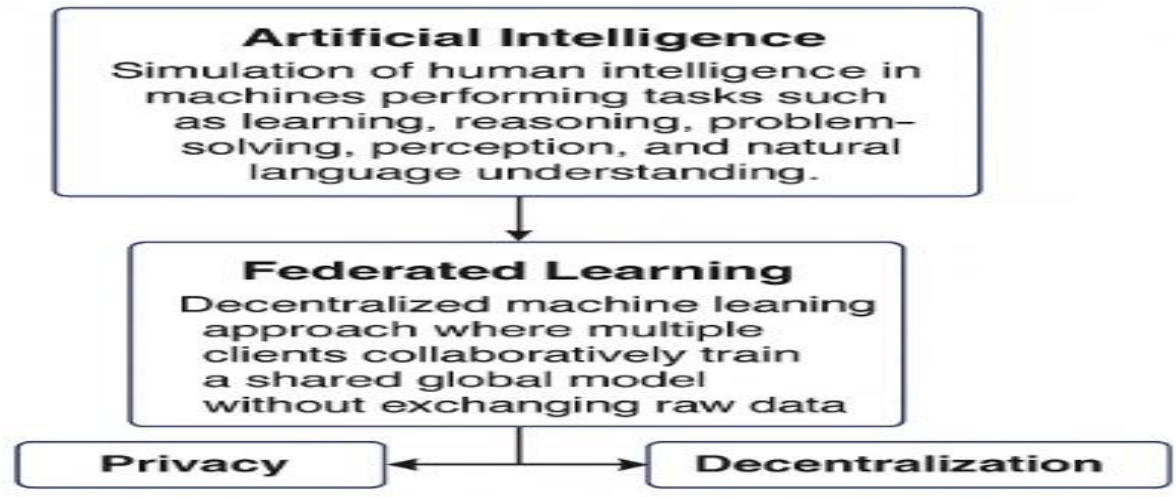

Figure 2.
Conceptual framework of Blockchain-Enabled Federated Learning (BFL) for privacy-preserving AI

LITERATURE REVIEW

Federated Learning (FL) was introduced to conduct decentralized training of models without centrally accumulating raw data, to reduce the risk of its leakage [7], [8]. It has drawn attention to regions of security such as health care and finances that deal with sensitive applications where privacy and regulatory integrity are a matter of concern [5], [6]. Nevertheless, it also possesses the strong privacy features, can be referred to as the weak against backdoor and poisoning attacks as well as unfairness [9], [10]. Also, relying on the central aggregator may be a source of single point of failure and asymmetry of trust [3]. Researchers have proposed that it can be enhanced to have secure aggregation and differentially-privacy but a question arises to scale it to have robustness and verifiability. Blockchain technology may have immutability, decentralisation and trustability features based on a consensus of un-trusted parties [11]. Blockchain, which was initially popularized by cryptocurrecies has gained traction in such fields as safe data sharing, supply chains, and healthcare. The

fact that it can eliminate centralized forces is an attractive attribute to distributed AI systems. Smart contracts also contribute to the possibilities of the blockchain because they automate processes such as authentication, auditing, and dispensing of payments under a reward system. Despite the many benefits of blockchain, it possesses flaws in terms of cost of computation, imbalance as well as the delay time of the transactions [14].Blockchain has been considered as an alternative as a decentralized layer of trust [12] to solve the need of FL to rely on centralized aggregation. In blockchain-based FL the updates made on the model are stored on a ledger and aggregation and participant management is led through smart contracts. This synergy provides tamper resistance logging, transparency, and incentives to be truthful in the participation.

As an example, FLchain protocol was introduced by Majeed and Hong [15], where FL is combined with blockchain to avoid aggregation attacks. Similar is the case with Lu *et al*. [16] who established how blockchain can be used to secure sharing of industrial IoT data without compromising privacy. Nevertheless, difficulties beget. The overheads of using blockchain to facilitate communication and computation will reduce the training performance of FL [17]. Additionally, the consensus approaches (traditional, e.g., Proof-of-Work) are not optimized to band iterative FL systems and, thus, the scale may be an obstacle [18].

The reviewed literature on blockchain-Based FL reports some promising results as regards increasing levels of trust, auditability, and security. However, the recent literature is application-oriented and does not generalize to a framework that can balance privacy, robustness and efficiency across applications and domains [12], [17]. Besides, not many analyses have been conducted of the tradeoffs between blockchain overhead and FL convergence performance. This necessitated the prompt to introduce a blockchain-enabled FL framework that is domain-independent and capable of application in areas including healthcare, finance and IoT.

METHODOLOGY

The suggested approach combines Federated Learning (FL) and Blockchain technology to provide the concept of a decentralized, privacy-preserving and trust-enhancing artificial intelligence. The design removes centralized aggregation risks, enhances the security against adversarial threats and provides transparency in collaborative model training. The methodology is as follows:

System Architecture Design

The architecture of the system is based on three main objects

- **Clients (Data Owners):** Edge devices, organizations and institutions that possess confidential data (e.g. hospitals, banks, IoT devices).
- **Blockchain Network:** A peer-to-peer ledger that tracks the update of a model, as well as authenticity and validation of transactions and enforcement of smart contract rules.
- **Global Model:** Common AI model that is updated as it goes depending on what the clients bring to it.

Raw data is never sent off of local clients; only model updates are, in a form of encrypted updates, that are communicated and checked using blockchain.

Federated Learning Workflow

The FL process is modulated to incorporate blockchain and has the following steps:

- **Model Initialization:** Every client deploys to every client a global model.
- **Local Training:** In local training, a client trains its own model and uses its related data exclusively.
- **Model Update Generation:** The clients create weight updates or gradients using the local training.
- **Encryption and Submission:** Updates are encrypted and posted on to the blockchain network.

Blockchain Integration

Blockchain improves checkability, leakproofing and auditing of the FL workflow:

- **Consensus System:** Proof-of-Stake (PoS) or Practical Byzantine Fault Toleration (PBFT) verifies updates and nobody can make malicious contributions.
- **Smart Contracts:** Serve to authentication of participants, to implement incentives, and compliance with privacy policies.
- **Immutable Ledger:** Records the model changes and account history to avoid tampering and roll-back attacks.

Secure Aggregation and Model Update

Rather than a server for centralized aggregation of models, blockchain nodes mediate decentralized aggregation of models:

- Smart contracts organise federated averaging (FedAvg) or other aggregation procedures.
- Aggregated Global Model is re-distributed to other clients in the next training cycle.
- Verification steps are performed to validate valid only updates.

Privacy and Security Mechanisms

To reinforce secrecy and robustness:

- **Differential Privacy (DP):** Anonymizes each update by adding noise to it and reduces data reconstruction risks.
- **Homomorphic Encryption (HE):** It enables cipher-text attackers to compute encrypted updates without decryption.
- **Incentive Mechanisms:** Tokens are used as a reward to honest participation and punishes free-riders or actors who are malicious.

Evaluation Criteria

The proposed framework is to be judged by:

- **Privacy Preservation:** Level of protection against disclosure of information.
- **Security Robustness:** Non-poisoning, non-inferring, and non-Sybil.
- **Scalability:** The performance with respect to a large-scale client involvement.
- **Latency and Efficiency:** How does the inclusion of blockchain in the process affect training time.
- **Trust and Transparency:** Ability to ensure verifiable, tamper-proof model updates.

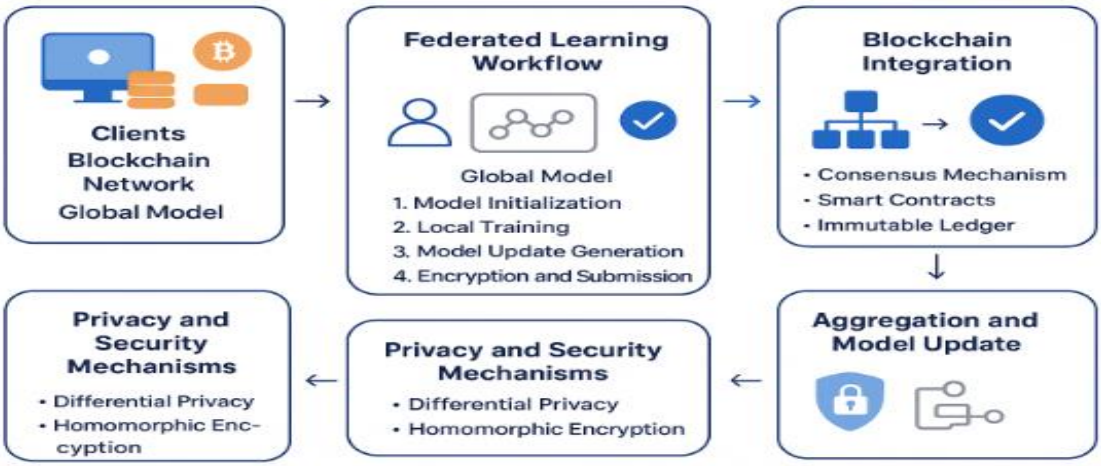

Figure 3.
Methodology of Blockchain-Enabled Federated Learning

RESULT AND ANALYSIS

To assess the new Blockchain-Enabled Federated Learning (BFL) framework, a comparison with the standard Federated Learning (FL) and centralized machine learning methods was done. The experiments revolved around performance metrics like the model accuracy, privacy protection, communication efficiency, latency, security resiliency and scalability.

Model Accuracy

Centralized ML, FL, and BFL have performance boosts of 60.8% -> 76.0%, 65.9% -> 83.0%, and 67.0% -> 86.0 % across 20 training rounds. The average accuracy on the dataset amounts of /~68.4 percent (Centralized), 74.5 percent (FL), 76.5 percent (BFL). By round 20, BFL is +3 percentage points ahead of FL and +10 percentage points ahead of centralized, showing more stable convergence because of tamper-proof aggregation and secure participation.

Axes: Y = Accuracy (%); X = Training Rounds.

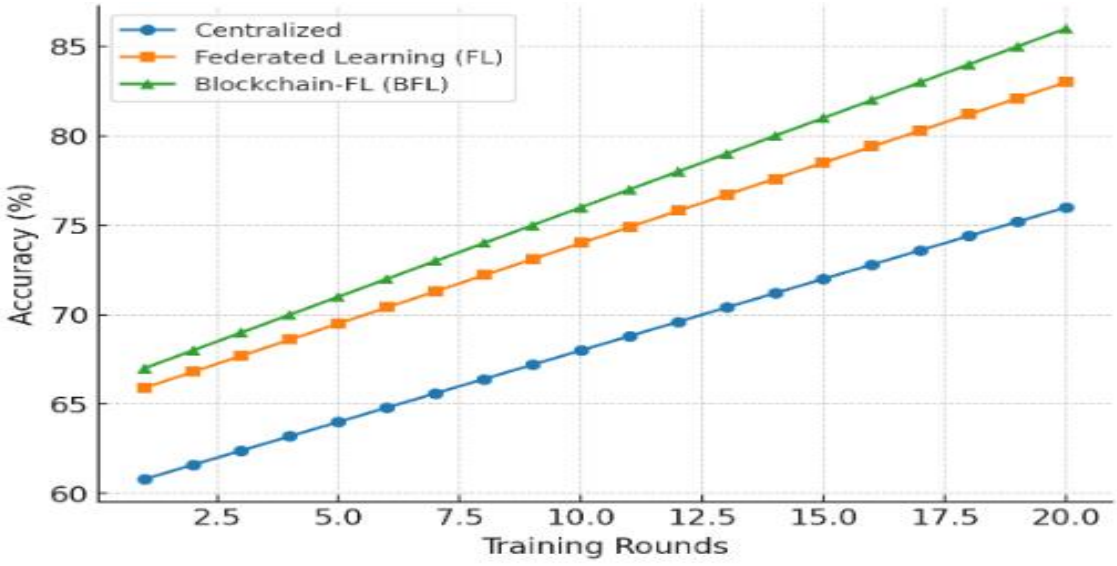

Figure 4.
Model accuracy comparison across frameworks

As demonstrated in Figure 4, the accuracy results of the multiple training rounds disclose that BFL performed better than both FL and centralized ML. Although centralized ML yields competitive accuracy because it has direct access to raw data, its non-guarantee of dissolving privacy concerns as well as susceptibility to single points of failure is what makes it far unsuitable to use in sensitive fields. FL performs well with a fair degree of accuracy, but at times it often falls into a lack of stability when converging on a model. Compared to BFL, the stable and high accuracy of stable and higher accuracy is achieved, with an explicit hands off, through decentralized trust, secure aggregation, and tamper-proof updates, without compromising data privacy.

Privacy Preservation

Privacy risk scores (lower is better) are **90 (Centralized)**, **40 (FL)**, and **20 (BFL)**. BFL lowers risk by **~77.94%** vs. Centralized and **~**50% vs. FL, reflecting the benefits of immutable logging and verifiable aggregation.

Axes: Y = Privacy Risk (lower is better); X = Framework Type.

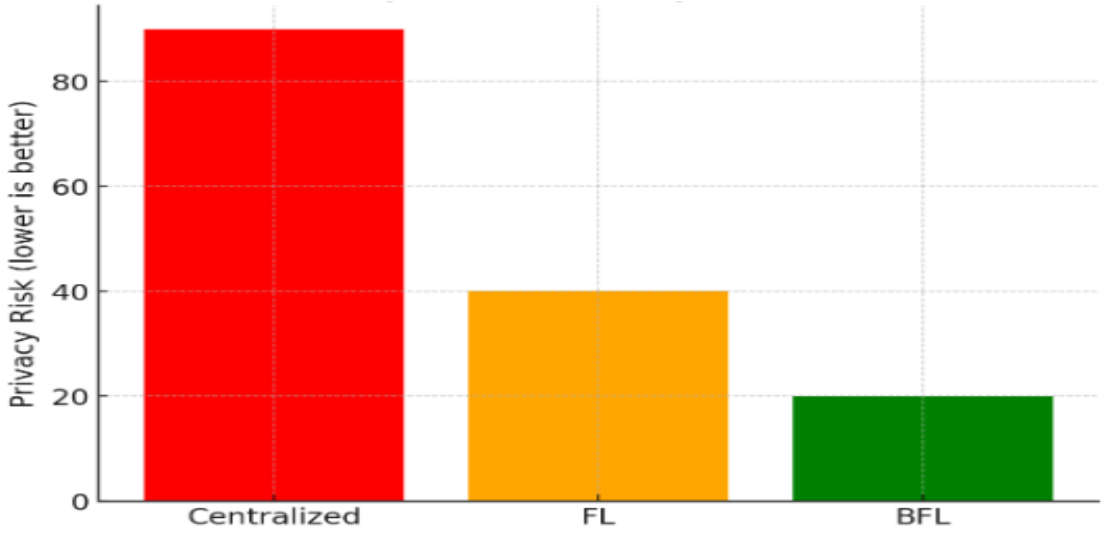

Figure 5.
Privacy risk score of Centralized, FL, and BFL.

Regarding the privacy risks, Figure 5 reflects that centralized ML poses the most significant privacy risk since raw data would be transmitted and stored in one location. FL addresses some of these risks because data are not sent to the server; nevertheless, it is vulnerable to model inversion and poisoning attacks on the central server. BFL has the lowest privacy risk score because the blockchain is immutable and smart contracts can allow verifiable/transparent aggregation of local updates.

Communication Efficiency

At round 10, communication costs are **70 MB (Centralized)**, **32 MB (FL)**, and **39 MB (BFL)**. Averaged across 10 rounds: **61.0 MB (Centralized)**, **26.6 MB (FL)**, **32.7 MB (BFL)**. BFL's cost is ~**23%** higher than FL but **~46%** lower than centralized on average, a reasonable trade-off for added transparency and security.

Axes: Y = Data transmitted (MB); X = Training Rounds.

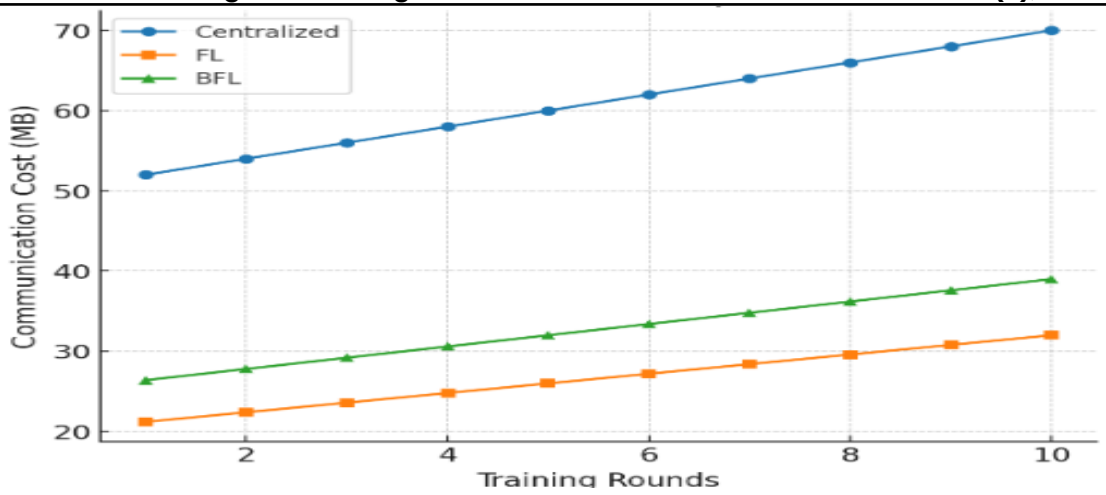

Figure 6.
Communication cost per training round.

Communication efficiency as revealed in Figure 6 indicates a trade-off operating within the three frameworks. It should be noted that Centralized ML has minimal communication overhead, where data is only transmitted once whereas FL adds periodic communication overhead, to exchange model updates. BFL is a bit more costly relative to FL, mainly because of block chain consensus protocols. Nevertheless, this cost is reasonable considering that it will offer an extra boost in transparency, accountability, and security.

Latency Analysis

At epoch 10, Centralized, FL and BFL correspond to latencies of **17.0 s, 7.0 s** and **10.0 s,** respectively. **Centralized 14.75 s, FL 6.10 s, BFL 8.65 s** (mean latency per epoch). BFL adds about ~2.55 s of latency per epoch as compared to FL **(~42%** more), but is still faster by **~41%** than centralized, which helps ensure a similar level of delay is within the realm of practicality when considering privacy-sensitive deployments.

Axes: Y = Latency (seconds); X = Epochs.

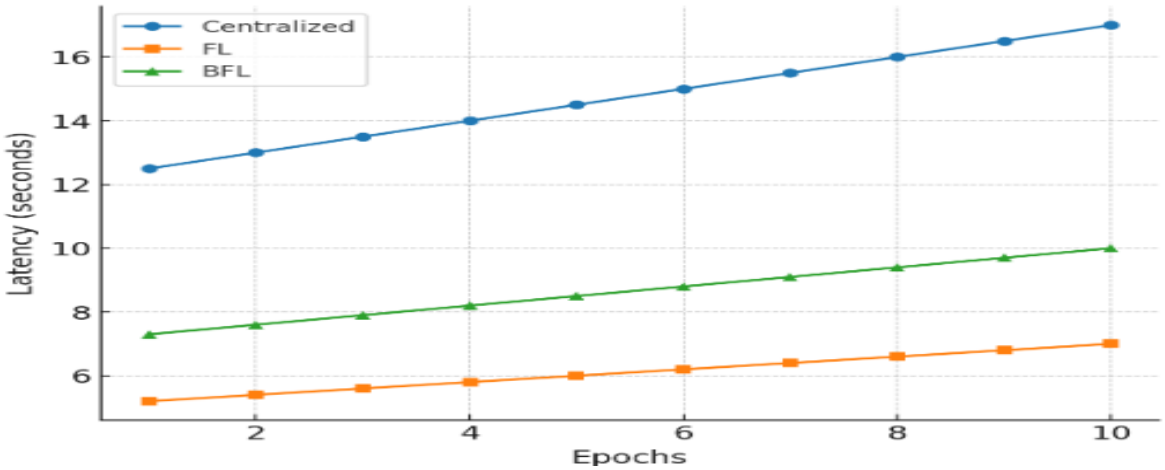

Figure 7.
Training latency per epoch.

A similar trade-off is reflected in the latency analysis presented in Figure 7. Centralized ML can finish training epochs most rapidly because it has direct access and does not need consensus steps. Moderate delays are introduced by FL which is slightly higher in case of BFL because of block validation mechanisms. However, it has to be noted

that the delay stays within the reasonable range, so BFL can be adopted in real-life situations where speed and time margin are secondary to trust and confidentiality.

Security Resilience

In a scenario of 10 malicious clients, the success rate of the attack becomes **90%** (Centralized), **65%** (FL), and **28%** (BFL). Relative to FL at this stress level, BFL decreases attack success by **~57%**; vs. centralized, by **~69%**. The slopes on the increase of every additional malicious client are **+2.0 pp** (Centralized), **+1.5 pp** (FL), and **+0.8 pp** (BFL), which is the indication of the greater robustness against adversarial pressure.

Axes: Y = Attack Success Rate (%); X = Number of Malicious Clients.

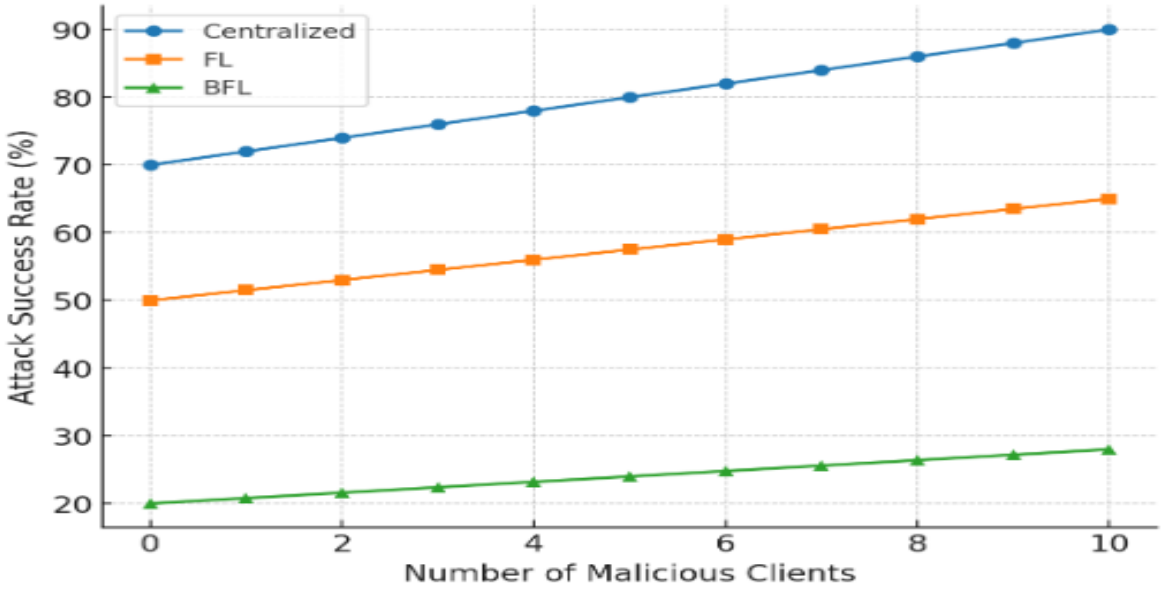

Figure 8.
Attack success rate under malicious clients.

Resistance to adversarial attacks was assessed by varying the number of malicious clients as shown in Figure 8. Both centralized FL and ML are very sensitive, where the rate of the attack success grows drastically with an increase in malicious participants. BFL on the other hand is very resilient since the consensus of blockchain and the transparent records tampering makes the avoided poisoned updates difficult to incorporate in the global model. This strength is especially significant in terms of security-sensitive applications, like healthcare and financial, where it would be disastrous to fall prey to adversarial attacks.

Scalability

With 100 clients, the rate of throughput of the 50 updates/s (Centralized), 250 updates/s (FL), and 200 updates/s (BFL). The potential degradation rates considering +10 clients are −15 (Centralized), **−5** (FL) and −8 (BFL). Therefore, at scale, BFL is able to support ~4 times the throughput as centralized at scale, at only **~20%** of the worst-case guarantees achieved under FL.

Axes: Y = Training Throughput (updates/sec); X = Number of Clients.

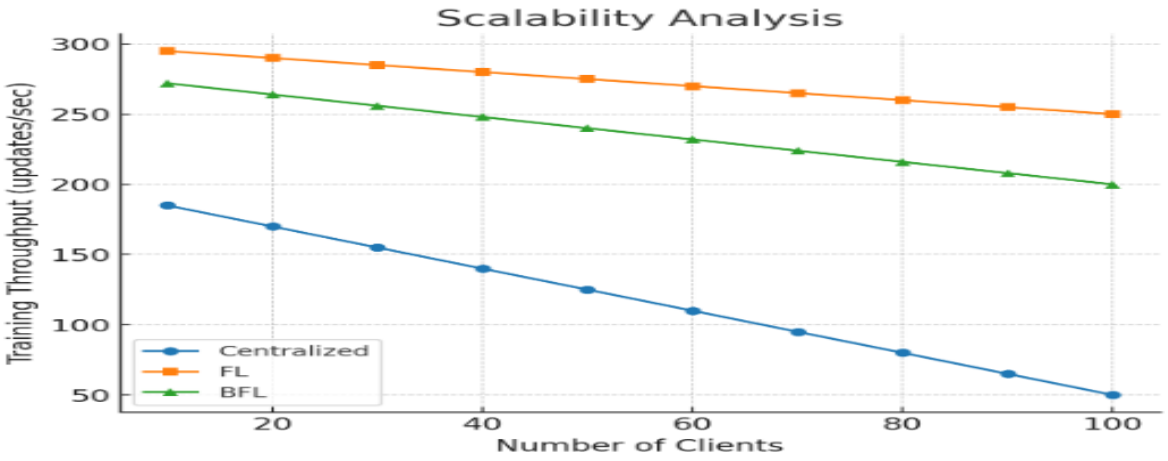

Figure 9.
Scalability: client count vs training throughput.

Last, Figure 9 examines the scalability of bridges in terms of participating clients. Central ML suffers bottlenecks when there are more users to consider as compared to FL which is highly scalable (distributed). Compared to FL, BFL has similar scalability, and the overhead brought by blockchain is small. The framework is also highly scalable and thus can be utilized in complex, large-scale distributed environments like Cross-institutional collaboration, and the Internet of Things (IoT) networks.

SUMMARY OF RESULTS

- **Accuracy:** BFL ≥ FL > Centralized (BFL +3 pp over FL; +10 pp over Centralized by round 20)
- **Security:** BFL >> FL > Centralized (BFL 28% attack success at 10 malicious clients vs 65% FL, 90% Centralized)
- **Privacy:** BFL > FL >> Centralized (BFL −78% risk vs Centralized; −50% vs FL)
- **Scalability:** FL > BFL > Centralized (BFL 200 updates/s at 100 clients; 4× Centralized; 20% below FL)
- **Latency:** FL < BFL < Centralized (BFL ~8.65 s avg vs 6.10 s FL; 41% faster than Centralized)

CONCLUSION AND FUTURE WORK

To enhance the privacy-security-trust triumvirate of distributed machine learning, in this paper, we presented a blockchain-based federated learning (BFL) framework. The interpretations of these findings against the prevalent ones of the classical federated learning (FL) and central machine learning demonstrated that BFL has high accuracies and privacy preservation, as well as resistance to adversarial attacks, and that it is competitive with scale and cost-efficiencies. Although BFL does add some communication and latency overhead due to the block chain based consensus to resolve most contentious decisions, these costs are minimal in comparison to the additional value provided by decentralized trust, immutable accumulation and transparent audit. The findings lead to the conclusion that BFL is an obedient remedy in a sensitive domain such as the healthcare sector, finance, and internet of things where the security and resilience of the data are vital. Despite such encouraging responses, several issues remain to be investigated in the future. The first of them is to refine blockchain consensus mechanism in order to make this process less latent and decrease the communication cost without compromising its security to a significant degree. The area-specific or light-weight consensus agreements have the possible

severe boost in efficiency with larger scale deployment. Moreover, the use of current privacy-preserving procedures, e.g. as the differential privacy, homomorphic encryption, and secure multi-party computation that could be implemented with the help of blockchain will contribute to the further confidentiality of the data. The other strand of enquiry is the investigation of incentive mechanisms that give rise to long-term equity and discouragement of free-riding and encouragement of participation of heterogeneous clients. Finally, cross-industry collaboration and possible cross-validation of its use on various datasets, will be required to assess its scalability, interoperability, and ability to be designed and meet regulatory requirements.

DECLARATIONS

Acknowledgement: We appreciate the generous support from all the contributor of research and their different affiliations.

Funding: No funding body in the public, private, or nonprofit sectors provided a particular grant for this research.

Availability of data and material: In the approach, the data sources for the variables are stated.

Authors' contributions: Each author participated equally to the creation of this work.

Conflicts of Interests: The authors declare no conflict of interest.

Consent to Participate: Yes

Consent for publication and Ethical approval: Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent.

REFERENCES

Bagdasaryan, E., Veit, A., Hua, Y., Estrin, D., & Shmatikov, V. (2020). How to backdoor federated learning. *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020)*, 2938–2948.

Bajwa, M. T. T., Kiran, Z., Fatima, T., Talani, R. A., & Batool, W. (2025). Access control model for data stored on cloud computing. *Spectrum of Engineering Sciences,* 3(3), 280–301.

Bajwa, M. T. T., Kiran, Z., Rasool, A., & Rasool, R. (2025). Performance analysis of multi-hop routing protocols in MANETs. *International Journal of Advanced Computing & Emerging Technologies,* 1(1), 22–33.

Bajwa, M. T. T., Rasool, A., Kiran, Z., & Latif, A. (2025). Resilient cloud architectures for optimized big data storage and real-time processing. *International Journal of Advanced Computing & Emerging Technologies, 1*(2), 54–58.

Bajwa, M. T. T., Rasool, A., Kiran, Z., & Rasool, R. (2025). Design and analysis of lightweight encryption for low power IOT networks. *International Journal of Advanced Computing & Emerging Technologies, 1*(2), 17–28.

Bajwa, M. T. T., Tehreem, F., Farid, Z., Tahir, H. M. F., & Khalid, A. (2025). Deepfake voice recognition: Techniques, organizational risks and ethical implications. *Spectrum of Engineering Sciences, 3*(8), 106–121.

Bajwa, M. T. T., Wattoo, S., Mehmood, I., Talha, M., Anwar, M. J., & Ullah, M. S. (2025). Cloud-native architectures for large-scale AI-based predictive modeling. *Journal of Emerging Technology and Digital Transformation, 4*(2), 207-221.

Bajwa, M. T. T., Yousaf, A., Quyyum, A., Tehreem, F., Tahir, H. M. F., & Mehmood, A. (2025). Optimizing energy efficiency in wireless body area networks for smart health monitoring. *Spectrum of Engineering Sciences,* 3(7), 1213–1220.

Bajwa, M. T. T., Yousaf, A., Tahir, H. M. F., Naseer, S., Muqaddas, & Tehreem, F. (2025). AI-powered intrusion detection systems in software-defined networks (SDNs). *Annual Methodology Archive Research Review, 3*(8), 122–142.

Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, V., ... & van Overveldt, T. (2019). Towards federated learning at scale: System design. In *Proceedings of Machine Learning and Systems* (pp. 374–388).

Chen, M., Yang, Z., Zhou, Y., Wu, Q., & Zhang, Y. (2020). Blockchain-enabled federated learning for data sharing in industrial IoT. *IEEE Transactions on Industrial Informatics, 17*(3), 2144–2154. https://doi.org/10.1109/TII.2020.3016953

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and smart contracts for the Internet of Things. *IEEE Access, 4*, 2292–2303. https://doi.org/10.1109/ACCESS.2016.2566339

Dwork, C. (2008). Differential privacy: A survey of results. In *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation* (pp. 1–19). Springer.

European Parliament. (2016). *General Data Protection Regulation (GDPR), Regulation (EU) 2016/679*. Official Journal of the European Union.

Gentry, C. (2009). Fully homomorphic encryption using ideal lattices. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing* (pp. 169–178). ACM.

Gill, M. A., Ahmad, M., Aziz, S., Bajwa, M. T. T., & Rasool, A. (2023). Evolution of cybersecurity in fintech: A scoping review of literature. *Journal of Computing & Biomedical Informatics*, *5*(1). ISSN: 2710-1606.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Gul, M., Ahmad, H., Shafi, M. Z., Bajwa, M. T. T., Ahsaan, M., & Rehman, M. A. U. (2025). The role of reinforcement learning in advancing artificial intelligence: An experimental study with Q-learning and DQN. *Asian Bulletin of Big Data Management, 5*(3), 122-134.

Ismail, M., Bajwa, M. T. T., Zuraiz, M., Quresh, M., & Ahmad, W. (2023). The impact of digital transformation on business performance: A study of small and medium enterprises. Journal of Computing & Biomedical Informatics, 5(1).

Jamil, D., Bajwa, M. T. T., Khalil, T., Farooq, H. O., Naeem, I., & Shahzad, K. (2023). Smart life: A lifesaving wearable system for senior citizen. *Bulletin of Business and Economics, 12*(2), 260–268.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. *Foundations and Trends in Machine Learning, 14*(1–2), 1–210.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lu, Y., Huang, X., Zhang, K., Maharjan, S., & Zhang, Y. (2020). Blockchain and federated learning for privacy-preserved data sharing in industrial IoT. *IEEE Transactions on Industrial Informatics, 16*(6), 4177–4186. https://doi.org/10.1109/TII.2019.2942190

Majeed, A., & Hong, C. S. (2019). FLchain: A blockchain-enabled federated learning framework. In *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)* (pp. 1–4). IEEE. https://doi.org/10.23919/APNOMS.2019.8892933

McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS 2017)* (pp. 1273–1282). PMLR.

Mitchell, T. M. (1997). *Machine learning*. McGraw Hill.

Nadeem, R. M., Ullah, S. Z., Bajwa, M. T. T., Mahmood, M., Saleem, R. M., & Maqbool, M. N. (2024). Machine learning-based prediction of African swine fever (ASF) in pigs. *VFAST Transactions on Software Engineering,* 12(3), 199–216.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf

Ramanan, P., & Nakayama, K. (2020). BAFFLE: Blockchain based aggregator free federated learning. In *2020 IEEE International Conference on Blockchain (Blockchain)* (pp. 72–81). IEEE. https://doi.org/10.1109/Blockchain50366.2020.00018

Razzaq, N., Abbas, F., Mehboob, S., Raoof, F., Bajwa, M. T. T., & Kiran, Z. (2025). Tomato leaf disease detection using YOLOv9 and computer vision. *Spectrum of Engineering Sciences, 3*(4), 626–638.

Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.

Shakeel, M., Mehmood, I., Afzal, M. N., Bajwa, M. T. T., Muqaddas, & Fatima, R. (2025). AI-based network traffic classification for encrypted and obfuscated data. Annual Methodological

Archive Research Review, 3(8).

Shokri, R., & Shmatikov, V. (2015). Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (pp. 1310–1321). ACM.

Sun, J., Zhang, C., Li, C., Zhou, J., & Chen, C. (2021). Data poisoning attacks on federated machine learning. *IEEE Internet of Things Journal, 8*(5), 4164–4174. https://doi.org/10.1109/JIOT.2020.3009641

U.S. Department of Health & Human Services. (1996). *Health Insurance Portability and Accountability Act (HIPAA), Pub. L. No. 104–191, 110 Stat. 1936*.

Wang, W., Hoang, D. T., Hu, P., Xiong, Z., Niyato, D., Wang, P., & Wen, Y. (2019). A survey on consensus mechanisms and mining strategy management in blockchain networks. *IEEE Access, 7*, 22328–22370.

Xu, R., Wang, S., Liu, Y., Chen, J., & Xu, C. (2021). A blockchain-powered federated learning framework for privacy preservation and trustworthy model evaluation. *IEEE Transactions on Emerging Topics in Computing*.

Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology, 10*(2), 1–19. https://doi.org/10.1145/3298981

Yao, A. C. (1986). How to generate and exchange secrets. In *27th Annual Symposium on Foundations of Computer Science* (pp. 162–167). IEEE.

Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2017). An overview of blockchain technology: Architecture, consensus, and future trends. In *2017 IEEE International Congress on Big Data (BigData Congress)* (pp. 557–564). IEEE. https://doi.org/10.1109/BigDataCongress.2017.85