Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

A Comprehensive Survey of Federated Learning Approaches for Privacy-Preserving Machine Learning

Research Article https://stem.techspherejournal.com

Author Details
Akinsiku Ayokunle Michael
Computer Science Department, The Federal Polytechnic Ado-Ekiti, Ekiti State, Nigeria.

****Corresponding author's email***: akinsiku_am@fedpolyado.edu.ng
DOI: https://doi.org/10.5281/zenodo.15830919

ABSTRACT

Federated Learning (FL) has emerged as a promising approach to privacy-preserving machine learning (PPML), allowing multiple clients to collaboratively train models without sharing their raw data. This paradigm addresses critical privacy, security, and regulatory concerns that hinder traditional centralized machine learning, especially in sensitive domains such as healthcare, finance, and edge computing. This paper presents a comprehensive survey of FL algorithms, examining classical methods like FedAvg and FedSGD, optimization-aware approaches such as FedProx and Scaffold, communication-efficient techniques, and privacy-enhanced frameworks integrating differential privacy, homomorphic encryption, and secure multiparty computation. Beyond algorithmic analysis, the paper explores real-world applications of FL across domains where data sensitivity and decentralization are paramount. It also discusses prevailing threats—gradient leakage, model inversion, poisoning, and backdoor attacks—and outlines corresponding mitigation strategies. Key challenges such as client heterogeneity, communication overhead, and trust assumptions are critically examined. The study identifies open research issues, including the need for scalable personalization, incentive mechanisms, integration with explainable AI, federated reinforcement learning, and deployment in low-resource environments. This survey provides a foundational understanding of federated learning's current landscape and future potential, offering insights for researchers and practitioners aiming to develop secure, efficient, and inclusive decentralized AI systems.

Keywords: Federated Learning, Privacy-Preserving Machine Learning, Decentralised AI, Security in Distributed Systems, Differential Privacy.

1 Introduction

The proliferation of machine learning (ML) techniques has revolutionized industries by enabling data-driven automation, prediction, and decision-making at an unprecedented scale [1]. From healthcare diagnostics to financial forecasting, and smart devices to autonomous vehicles, ML models have become central to digital transformation efforts. These models, however, traditionally rely on centralized training pipelines, where large volumes of data are aggregated from various sources and processed in a central server. While effective, this centralized paradigm raises significant concerns regarding data privacy, security, and ownership, particularly in sensitive domains such as healthcare, finance, and edge-based computing environments.

In recent years, the societal and regulatory landscape surrounding data privacy has evolved dramatically. Stringent data protection laws like the General Data Protection Regulation (GDPR) in Europe, the Health Insurance Portability and Accountability Act (HIPAA) in the United States, and similar frameworks globally have underscored the need for

privacy-preserving alternatives in machine learning [2]. These regulations emphasize user consent, data minimization, and secure processing, often making centralized data collection impractical or legally restricted. Furthermore, public awareness of data misuse and cyber threats has heightened expectations for transparency and ethical data usage.

FL has emerged as a promising solution to these challenges. Rather than moving data to a central server, FL shifts the model training process to the data source [3]. Data remains on local devices or organizational silos, and only model updates like gradients or weights are shared with a central aggregator. This decentralized learning paradigm preserves data locality, enhances data security, and promotes user autonomy. Importantly, it aligns well with privacy-focused requirements by reducing the risk of sensitive data exposure during training. Additionally, FL enables collaboration across institutions and devices without violating data ownership boundaries.

The motivation for federated learning stems from three interrelated drivers:

1. **Decentralization:** Distributing computation across edge devices or institutional silos allows for more resilient and scalable learning systems.
2. **Data Ownership:** FL respects the autonomy of data producers, be it hospitals, banks, or individuals, by ensuring data remains under local control.
3. **Regulatory Compliance:** By avoiding raw data transmission, FL offers a pathway to building compliant AI solutions in regulated industries.

Despite these benefits, implementing FL introduces its own set of challenges. Heterogeneous data distributions, limited device resources, communication bottlenecks, and threats such as poisoning attacks or gradient leakage present significant barriers [4]. These risks have driven the development of privacy-preserving mechanisms, including differential privacy, secure multiparty computation, and homomorphic encryption, designed to reinforce trust and security in federated systems.

1.1 Objective of the Paper

This paper aims to provide a comprehensive survey of federated learning approaches with a focus on privacy-preserving machine learning. Specifically, it seeks to:

a. Survey state-of-the-art federated learning algorithms, highlighting their principles, strengths, and limitations.
b. Examine the application of FL in real-world domains such as healthcare, finance, and edge computing, where data privacy is critical.
c. Investigate privacy and security challenges, along with the techniques developed to mitigate them.

2 Fundamentals of Federated Learning

2.1 Federated Learning

Federated Learning (FL) is a distributed machine learning paradigm that allows multiple clients like mobile devices, edge nodes, or institutions, to collaboratively train a shared global model without exposing their local data [5]. First introduced by Google in 2016, FL was conceived as a response to the growing concern over user privacy and data security in centralized machine learning systems. Instead of transferring raw data to a centralized server, FL brings the model to the data performing training locally on devices and only sharing model updates like gradients or weights, with a coordinating server.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

2.1.1 Architecture of Federated Learning

There are two main architectural approaches:

i. **Centralized Federated Learning:** A central server orchestrates the training by distributing the global model to participating clients and aggregating their local updates to improve the model iteratively. This is the most widely used configuration [6][7].

ii. **Decentralized Federated Learning:** There is no central coordinator. Instead, clients communicate directly with one another (peer-to-peer) to train the model. This model improves robustness but adds complexity to communication and convergence [7].

2.1.2 Types of Federated Learning

Federated learning can be categorized based on how data is distributed among clients:

i. **Horizontal Federated Learning (HFL):** Clients share the same feature space but differ in samples. For example, multiple hospitals with the same patient records structure but different patients [8].

ii. **Vertical Federated Learning (VFL):** Clients share the same sample space but have different features. For example, a bank and an e-commerce platform may serve the same customers but collect different types of data [9].

iii. **Federated Transfer Learning (FTL):** Clients differ in both sample and feature spaces but may still benefit from knowledge transfer via shared learning tasks. FTL is especially useful in cross-domain collaboration scenarios with limited data overlap [10].

FL thus represents a significant shift from traditional centralized training, promoting decentralized intelligence and privacy-aware machine learning.

2.2 Key Components of Federated Learning

Federated Learning systems involve multiple technical components working in tandem to facilitate secure and efficient model training.

a. Client

Clients are the local data holders participating in training. They could be mobile devices, edge sensors, or organizational silos. Each client trains the model on its local data using predefined training epochs and then sends updates to the aggregator [11].

b. Server (Aggregator)

The server is responsible for orchestrating the overall training process. In centralized FL, it broadcasts the initial global model to the clients and aggregates the local updates (often via weighted averaging, like FedAvg) to form a new global model [12]. This process is repeated iteratively.

c. Communication Protocols

Communication in FL systems must be efficient and secure, as it often takes place over bandwidth-limited or untrusted networks [7]. Protocols manage:

i. Transmission of model updates.
ii. Client-server synchronization.
iii. Encryption and privacy-preserving mechanisms.

d. Training and Aggregation Phases

1. **Training Phase:** Clients train the model locally using their private datasets and compute local gradients or parameter updates.
2. **Aggregation Phase:** The server collects updates from all (or a selected subset of) clients and combines them into a new global model.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

The process is repeated over multiple communication rounds until model convergence is achieved or a performance threshold is met [7].

2.3 Advantages over Traditional Machine Learning

Federated Learning provides several advantages that make it highly attractive, especially in privacy-sensitive and distributed environments:

a. **On-Device Learning**
FL enables training directly on user devices such as smartphones or IoT nodes, allowing intelligence to be embedded at the edge. This reduces reliance on cloud infrastructure and enhances user experience through personalized models [13].

b. **Data Privacy Preservation**
By keeping data local, FL minimizes the exposure of sensitive information and mitigates the risk of data breaches, eavesdropping, or unauthorized access during transmission. It is inherently more privacy-preserving than centralized learning, where data must be uploaded to a server [8].

c. **Reduced Latency and Bandwidth Usage**
Transmitting model updates instead of raw data conserves bandwidth, especially useful in mobile or IoT settings. Additionally, on-device processing allows for quicker inference and adaptation to local conditions [14].

d. **Regulatory Compliance**
FL aligns with legal frameworks such as the General Data Protection Regulation (GDPR) and Health Insurance Portability and Accountability Act (HIPAA) [15]. It offers a viable path to machine learning in regulated sectors by minimizing data movement and centralization.

3 Federated Learning Algorithms

Federated Learning has sparked a wave of innovation in distributed machine learning, leading to the development of diverse algorithms designed to optimize learning under different constraints such as data heterogeneity, communication bottlenecks, and privacy preservation. This section categorizes and explains key FL algorithms across four main dimensions: classical methods, optimization-aware techniques, communication-efficient strategies, and privacy-enhancing protocols.

3.1 Classical Approaches

FedAvg is the foundational algorithm in federated learning. It operates by sending the global model to a subset of clients, where each performs several local updates using stochastic gradient descent (SGD) on its private data [16]. These local models are then averaged (weighted by the number of samples) by the central server to update the global model.

a. **Performance:** FedAvg works efficiently in scenarios with independent and identically distributed (IID) data and moderate device heterogeneity. It significantly reduces communication rounds compared to fully synchronous distributed SGD.

b. **Limitations:** In non-IID data settings (which are typical in FL), FedAvg suffers from weight divergence, poor generalization, and fairness issues. Additionally, client dropout and limited compute power impact convergence.

FedSGD is a simpler alternative to FedAvg, where each participating client computes gradients on its local batch and sends them to the server at every iteration without performing multiple local epochs [17]. The server then averages these gradients to update the global model.

a. Strengths: FedSGD ensures tighter synchronization and is closer to centralized SGD.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

b. Limitations: High communication cost due to more frequent updates and limited client-side computation makes it inefficient in edge scenarios

3.2 Optimization-Aware Methods

To address non-IID data, device heterogeneity, and convergence inefficiencies, a range of optimization-aware FL algorithms have been proposed.

FedProx extends FedAvg by introducing a proximal term in the client's local objective function. This penalizes updates that deviate significantly from the global model and hence stabilizes training on heterogeneous devices [18].

a. Advantage: Handles system and statistical heterogeneity by preventing clients from drifting too far during local updates.
b. Limitation: May lead to slower convergence due to the regularization constraint.

Scaffold mitigates client drift caused by non-IID data through the use of control variates that track update directions and correct local model deviations.

a. Advantage: Faster convergence and better performance in non-IID settings.
b. Limitation: Requires storage and synchronization of control variates, increasing complexity.

FedNova normalizes local updates to ensure that the contribution of each client to the global model is independent of local update steps or batch sizes [19].

a. Advantage: Fairer contribution from clients and improved performance on unbalanced datasets.
b. Limitation: Slightly more complex update logic.

These optimization-aware algorithms improve training stability and generalization in real-world settings, where client behaviour and data are highly variable.

3.3 Privacy-Enhanced Algorithms

Despite keeping data local, FL is not inherently immune to privacy threats. Adversaries can reconstruct sensitive information from gradients or infer membership in the training data. To mitigate these risks, several privacy-preserving methods have been integrated into FL pipelines:

Differential Privacy (DP)

Adds random noise to local model updates before sharing with the server, ensuring that individual data points cannot be inferred. Key parameters include the privacy budget (ε) and noise scale [20].

- Trade-off: Stronger privacy implies reduced model utility.

Homomorphic Encryption (HE)

Enables computations (e.g., addition, multiplication) on encrypted data. Clients send encrypted updates, and the server aggregates them without decryption [21].

- Strength: Ensures confidentiality of updates.
- Limitation: Computationally expensive, especially for deep models.

Secure Multi-party Computation (SMPC)

Distributes the computation across multiple parties so that no single party learns any sensitive information. Used for secure aggregation [22].

- Advantage: Removes the need to trust a single aggregator.
- Complexity: High protocol overhead and latency.

These mechanisms can be integrated with standard FL algorithms to create robust privacy-preserving federated learning (PPFL) pipelines.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

3.4 Comparative Summary of Algorithms

Table 1 shows the comparative summary of the algorithms in FL.

Table 1: Comparison of Algorithms

Algorithm	Main Idea	Strengths	Limitations
FedAvg	Local model averaging	Simple, scalable	Poor performance on non-IID data
FedSGD	Synchronized gradient aggregation	Easier convergence analysis	High communication cost
FedProx	Adds regularization term	Handles heterogeneity	Slower convergence
Scaffold	Uses control variates to reduce drift	Improves convergence with non-IID data	Requires additional computation and storage
FedNova	Normalized update contribution	Fairness, better handling of unbalanced data	Slightly more complex to implement
DP-FL	Adds noise to updates	Differential privacy guarantee	Accuracy-privacy trade-off
HE-FL	Encrypts updates before aggregation	Strong data confidentiality	High computational overhead
SMPC-FL	Distributed secure computation	No need for a trusted server	High latency and protocol complexity

4 Application Domains of Federated Learning

Federated Learning has gained traction across multiple domains where data privacy, real-time processing, and regulatory compliance are paramount. By enabling decentralized model training, FL makes it possible to develop robust AI systems while preserving the confidentiality of sensitive data. This section explores three primary domains where FL has seen significant application: healthcare, finance, and edge computing with IoT devices.

4.1 Healthcare

The healthcare sector is one of the most critical beneficiaries of federated learning, given the sensitivity of patient data and the legal implications of data sharing. Traditional approaches that require centralizing health records across hospitals or clinics pose privacy risks and often breach data protection regulations like HIPAA and GDPR [23]. FL addresses these issues by allowing hospitals, laboratories, and wearable device manufacturers to collaboratively train models without sharing raw patient data.

Use Cases

a. **Collaborative Disease Prediction:** Multiple hospitals can jointly develop models for predicting diseases such as cancer, Alzheimer's, or COVID-19, using localized patient data.
b. **Wearable Health Monitoring:** Smart devices such as fitness trackers and ECG monitors can locally train models to predict arrhythmia, stress levels, or sleep disorders, with updates securely aggregated in the cloud.

Key Studies

a. **Google's Gboard (initial demonstration of FL):** Although not directly healthcare-focused, it set the precedent for privacy-conscious learning on user devices [24].
b. NVIDIA Clara: A FL platform tailored for medical imaging and diagnostics, enabling institutions to train models for radiology and pathology without exposing medical data.

Benefits

a. Protects patient confidentiality and complies with health data regulations.
b. Promotes collaboration across medical institutions for data-rich, generalized models.
c. Facilitates on-device intelligence in wearable health monitors.

Challenges

a. Data heterogeneity due to varied medical equipment, coding standards, and patient demographics.
b. Limited compute resources in hospital systems or edge health devices.

4.2 Finance

Financial institutions deal with vast quantities of highly sensitive data involving customer transactions, credit histories, and behavioural analytics. The potential of FL in this domain lies in enabling multiple institutions (banks, insurance companies, fintech platforms) to build powerful machine learning models without exposing proprietary or personally identifiable information (PII) [13].

Use Cases

a. Fraud Detection: FL allows institutions to detect fraud by learning patterns across multiple organizations while maintaining data confidentiality.
b. Credit Risk Assessment: Multiple banks can collaboratively assess creditworthiness using distributed financial records, leading to more robust risk models.
c. Personalized Banking: FL enables on-device learning for personal finance apps, providing customized insights without sending data to the cloud.

Key Benefits

a. Regulatory compliance with data protection laws (e.g., GDPR, PCI-DSS).
b. Cross-institutional intelligence without compromising competitive or private information.
c. Enhanced fraud detection through aggregated behavioural insights.

Challenges

a. Communication overhead due to high-frequency financial updates and latency-sensitive models.
b. Data imbalance and varying regulatory obligations across institutions.

4.3 Edge Computing and IoT

Edge computing and the Internet of Things (IoT) represent an ecosystem of interconnected devices, ranging from smart thermostats and wearables to autonomous vehicles and industrial sensors, that collect, process, and transmit data [5]. The nature of these devices demands learning models that can operate with minimal latency, bandwidth, and power.

Use Cases

a. Smart Homes: Devices such as smart speakers and cameras can learn user preferences locally, enabling better personalization without exposing private behaviors.
b. Autonomous Vehicles: FL enables vehicles to share learned models about road conditions or hazards without transmitting video or sensor data.
c. Mobile Devices: Smartphones can personalize voice assistants, keyboards (e.g., Gboard), and recommendation systems via on-device training.

Key Benefits

a. Enhances on-device learning and personalization.
b. Reduces dependency on cloud infrastructure, improving responsiveness.
c. Protects user privacy in sensitive, real-time environments.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

Challenges

a. Limited processing power, memory, and battery on edge devices.
b. Network instability and device availability issues hinder synchronous updates.
c. Security vulnerabilities in physical devices increase attack surfaces

4.4 Comparative Summary Table

Table 2 presents the comparative summary of the applications of FL

Table 2: Comparative Summary of the Applications of Federated Learning

Domain	FL Use Case	Key Benefits	Challenges
Healthcare	Diagnosis prediction, wearable health monitoring	Patient privacy, data compliance, cross-institution collaboration	Data heterogeneity, limited local compute power
Finance	Fraud detection, credit risk modelling, personalized banking	Regulatory compliance, data confidentiality, enhanced risk assessment	Communication overhead, data imbalance
Edge/IoT	Smart devices, autonomous vehicles, personalized assistants	On-device learning, low latency, reduced cloud reliance	Resource constraints, asynchronous updates

By enabling collaboration across isolated data silos, FL fosters innovation while upholding privacy, trust, and compliance. Its application in these sectors has not only demonstrated feasibility but also uncovered pressing challenges that guide ongoing research in model optimization, personalization, and secure aggregation protocols.

5 Privacy and Security Challenges in Federated Learning

While FL offers inherent privacy benefits by retaining data on local devices, it is not immune to threats. Exchanging model updates instead of raw data does reduce direct exposure, but recent studies show that gradients and parameters can still leak sensitive information. Moreover, FL systems are vulnerable to adversarial manipulation, which can undermine model accuracy, trustworthiness, and fairness [25]. This section explores key privacy threats, security risks, mitigation strategies, and the trade-offs involved in implementing PPFL.

5.1 Privacy Threats

Gradient Leakage

One of the most well-documented threats in FL is gradient leakage, where adversaries can reconstruct the original training data from shared gradients [26]. For instance, a malicious server or client observing the gradients over several rounds may infer private attributes such as facial images or health records.

Membership Inference Attacks

In this attack, adversaries attempt to determine whether a specific data point was part of a client's training set. This is especially concerning in healthcare and finance, where even the presence of a data point (e.g., a disease label) can be sensitive.

Tech-Sphere Journal of Pure and Applied Sciences (TSJPAS)
A Subsidiary of Tech-Sphere Multidisciplinary International Journal (TSMIJ)
Akinsiku. Vol 2, Issue 1, 2025 Publication Edition
ISSN: 3092-9598

Model Inversion Attacks

Adversaries exploit access to the model parameters or outputs to infer latent features of the training data. For example, model inversion can be used to reconstruct a blurred or abstract version of a patient's image or financial behaviour [27].

Cross-Silo vs. Cross-Device Threat Models

a. Cross-Silo FL involves a small number of reliable participants (e.g., hospitals, banks). The trust level is higher, but the impact of targeted attacks can be severe.
b. Cross-Device FL includes millions of unreliable and potentially compromised edge devices. Here, threat detection and mitigation are more challenging due to scale and heterogeneity.

5.2 Security Risks

Federated learning systems are also susceptible to various security attacks that aim to degrade model integrity or introduce malicious behaviours.

1. **Poisoning Attacks**
   a. **Data Poisoning:** Malicious clients introduce false labels or corrupted data during local training to mislead the global model [28].
   b. **Model Poisoning**: Instead of altering local data, adversaries directly manipulate model updates (e.g., by scaling gradients or injecting noise) before sending them to the server [29].

2. **Backdoor Attacks**
   Attackers embed hidden malicious behaviours into the global model that are only triggered under specific conditions like particular input patterns [30]. These backdoors can be introduced by compromised clients acting stealthily over time.

3. **Byzantine Clients**
   In a Byzantine setting, some clients may act arbitrarily or maliciously, sending invalid or corrupted updates that disrupt global model convergence [28].

4. **Adversarial Updates**
   Attackers may send adversarially crafted updates that subtly shift the global model's decision boundary, leading to misclassifications in targeted classes or inputs [31].

5.3 Trade-offs

Designing a federated learning system that balances performance, efficiency, and privacy is inherently complex. Table 3 show the key trade-offs for FL.:

Table 3: Trade-off for FL

Trade-off	Description
Accuracy vs. Privacy	Stronger privacy mechanisms (e.g., higher DP noise) typically degrade accuracy.
Computation vs. Communication	Encryption and privacy-preserving aggregation increase processing time and bandwidth requirements.
Security vs. Scalability	Techniques like SMPC and TEE enhance security but may not scale well in large deployments.

Responsiveness vs. Robustness	Asynchronous FL improves speed but complicates threat detection and correction.

Effective federated learning designs must make context-specific decisions on which trade-offs to prioritize balancing regulatory requirements, device capabilities, model performance, and user expectations.

6 Open Research Issues and Future Directions

As federated learning (FL) continues to gain adoption across diverse sectors, several unresolved challenges and emerging research opportunities warrant attention. These issues span algorithmic, infrastructural, ethical, and societal dimensions. Addressing them is essential to realizing the full potential of FL as a privacy-preserving machine learning paradigm. This section outlines key open research problems and promising directions for future work.

Research Directions:

a. Development of scalable federated optimization algorithms that support asynchronous updates, dynamic client selection, and adaptive learning rates.
b. Federated multi-task learning and meta-learning approaches that tailor global models to individual user contexts while maintaining shared knowledge.

6.1 Integrating Federated Learning with Explainable AI (XAI)

The black-box nature of many machine learning models poses a barrier to trust and transparency, particularly in high-stakes domains like healthcare and finance. Integrating FL with explainable artificial intelligence (XAI) is a critical research frontier [32].

Research Directions:

a. Designing interpretable federated models that provide local and global explanations without compromising privacy.
b. Developing methods for federated feature attribution, saliency mapping, and counterfactual reasoning across distributed clients.

6.2 Incentive Mechanisms for Participation

FL relies on voluntary participation by clients who may incur costs in terms of computation, battery consumption, and bandwidth. However, without proper incentive mechanisms, client engagement may be unsustainable.

Research Directions:

a. Application of game-theoretic and blockchain-based incentive schemes to reward honest participation.
b. Designing fair contribution evaluation metrics to allocate rewards based on data utility, update quality, or model performance impact.

6.3 Federated Reinforcement Learning (FRL)

Extending FL to support reinforcement learning (RL) is a nascent but exciting direction. In FRL, agents learn from distributed environments like smart vehicles or robots, while preserving privacy [33].

Research Directions:

i. Communication-efficient policy sharing and aggregation methods.
ii. Dealing with delayed feedback and non-stationary environments in distributed RL settings.

6.4 Benchmarking and Standardized Evaluation

Currently, there is a lack of standardized benchmarks and evaluation protocols for federated learning. Most FL research is evaluated using inconsistent datasets, metrics, and hardware configurations.

Research Directions:

i. Development of open-source FL simulation platforms like LEAF, Flower, FedML, with standardized datasets and metrics.
ii. Creating performance benchmarks that account for privacy guarantees, communication costs, fairness, and convergence behaviour.

6.5 FL in Low-Resource and Under-Represented Environments

Most FL studies are conducted in high-resource settings with stable connectivity and powerful hardware. However, its transformative potential lies in low-resource regions like rural healthcare, informal economies, developing countries, where data scarcity and privacy concerns are profound.

Research Directions:

i. Building lightweight, resource-efficient FL models compatible with low-power devices.
ii. Ensuring inclusivity and fairness by designing algorithms that mitigate bias against underrepresented groups and languages.

In summary, while federated learning has made significant progress, these open research questions highlight the need for interdisciplinary collaboration involving machine learning, systems engineering, security, economics, and ethics. Future advancements must focus not only on technical performance but also on societal relevance, inclusivity, and sustainability.

7 Conclusion

Federated Learning (FL) has emerged as a transformative paradigm for privacy-preserving machine learning (PPML), enabling collaborative model training without the need to centralize sensitive data. In an era increasingly governed by stringent data protection regulations and heightened public awareness of digital privacy, FL provides a principled approach to decentralizing intelligence while safeguarding individual and institutional data ownership. This paper presented a comprehensive survey of federated learning algorithms, highlighting classical methods such as FedAvg and FedSGD, optimization-aware techniques like FedProx and Scaffold, as well as communication-efficient and privacy-enhancing protocols involving differential privacy, homomorphic encryption, and secure multi-party computation. The practical applications of these algorithms were examined across key sectors—healthcare, finance, and edge computing/IoT, where privacy, latency, and personalization are critical to the success of AI systems. Despite its promising advantages, FL is not without challenges. The paper examined major privacy threats such as gradient leakage and model inversion, along with security risks including poisoning attacks and backdoors. Several technical countermeasures were discussed, though these often introduce trade-offs in accuracy, communication cost, and system complexity. The future of federated learning is rich with research opportunities—from scalability and personalization to federated reinforcement learning, incentive mechanisms, and the integration of explainable AI. Addressing these open questions is crucial for the responsible and robust deployment of FL at scale. In conclusion, federated learning holds significant promise as the backbone of decentralized, privacy-respecting artificial intelligence. Its ability to bridge collaboration and confidentiality makes it a compelling solution for building trustworthy AI in domains where data sensitivity, regulation, and real-world complexity converge. As research advances and tools mature, FL is poised to play a central role in shaping the next generation of ethical and secure AI systems.

.

References

[1] J. Ramya, S. S. Yerraguravagari, S. Gaikwad, and R. K. Gupta, "AI and Machine Learning in Predictive Analytics: Revolutionizing Business Strategies through Big Data Insights.," *Libr. Progress-Library Sci. Inf. Technol. Comput.*, vol. 44, no. 3, 2024.
[2] T. Bashir, "Zero Trust Architecture: Enhancing Cybersecurity in Enterprise Networks," no. September, 2024, doi: 10.32996/jcsts.
[3] A. Alzu'Bi, A. Alomar, S. Alkhaza'Leh, A. Abuarqoub, and M. Hammoudeh, "A review of privacy and security of edge computing in smart healthcare systems: issues, challenges, and research directions," *Tsinghua Sci. Technol.*, vol. 29, no. 4, pp. 1152–1180, 2024.
[4] A. Mora, A. Bujari, and P. Bellavista, "Enhancing generalization in federated learning with heterogeneous data: A comparative literature review," *Futur. Gener. Comput. Syst.*, 2024.
[5] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for internet of things: A comprehensive survey," *IEEE Commun. Surv. Tutorials*, vol. 23, no. 3, pp. 1622–1658, 2021.
[6] H. U. Manzoor, A. Shabbir, A. Chen, D. Flynn, and A. Zoha, "A survey of security strategies in federated learning: Defending models, data, and privacy," *Futur. Internet*, vol. 16, no. 10, p. 374, 2024.
[7] E. T. M. Beltrán *et al.*, "Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges," *IEEE Commun. Surv. Tutorials*, vol. 25, no. 4, pp. 2983–3013, 2023.
[8] C. Chen *et al.*, "Trustworthy federated learning: privacy, security, and beyond," *Knowl. Inf. Syst.*, vol. 67, no. 3, pp. 2321–2356, 2025.
[9] Y. Liu *et al.*, "Vertical federated learning: Concepts, advances, and challenges," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 7, pp. 3615–3634, 2024.
[10] I. Kevin, K. Wang, X. Zhou, W. Liang, Z. Yan, and J. She, "Federated transfer learning based cross-domain prediction for smart manufacturing," *IEEE Trans. Ind. Informatics*, vol. 18, no. 6, pp. 4088–4096, 2021.
[11] M. K. Kundalwal, A. Saraswat, I. Mishra, and D. Mishra, "Client Contribution Normalization for Enhanced Federated Learning," *arXiv Prepr. arXiv2411.06352*, 2024.
[12] M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, "Reviewing federated learning aggregation algorithms; strategies, contributions, limitations and future perspectives," *Electronics*, vol. 12, no. 10, p. 2287, 2023.
[13] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: Fundamentals, enabling technologies, and future applications," *Inf. Process. Manag.*, vol. 59, no. 6, p. 103061, 2022.
[14] A. Paju, M. O. Javed, J. Nurmi, J. Savimäki, B. McGillion, and B. B. Brumley, "Sok: A systematic review of tee usage for developing trusted applications," in *Proceedings of the 18th International Conference on Availability, Reliability and Security*, 2023, pp. 1–15.
[15] A. L. Imoize, M. S. Obaidat, and H. H. Song, "Legal implications of federated learning integration in digital healthcare systems," in *Federated Learning for Digital Healthcare Systems*, Elsevier, 2024, pp. 355–385.
[16] L. Collins, H. Hassani, A. Mokhtari, and S. Shakkottai, "Fedavg with fine tuning: Local updates lead to representation learning," *Adv. Neural Inf. Process. Syst.*, vol. 35, pp. 10572–10586, 2022.
[17] Y. Wang, S. Guo, D. Qiao, G. Liu, and M. Li, "FedSG: A Personalized Subgraph Federated Learning Framework on Multiple Non-IID Graphs," *IEEE Trans. Emerg. Top. Comput. Intell.*, 2024.
[18] J. Cui, Y. Li, Q. Zhang, Z. He, and S. Zhao, "A Federated Learning Framework Using FedProx Algorithm for Privacy-Preserving Palmprint Recognition," in *Chinese Conference on Biometric Recognition*, Springer, 2024, pp. 187–196.
[19] F. Li, X. Chen, K.-Y. Lam, B. Shen, and L. Wang, "Federated-learning-based wireless traffic prediction," in *2025 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, IEEE, 2025, pp. 748–753.
[20] A. El Ouadrhiri and A. Abdelhadi, "Differential privacy for deep and federated learning: A survey," *IEEE access*, vol. 10, pp. 22359–22380, 2022.
[21] Q. Xie *et al.*, "Efficiency optimization techniques in privacy-preserving federated learning with homomorphic encryption: A brief survey," *IEEE Internet Things J.*, vol. 11, no. 14, pp. 24569–24580, 2024.
[22] L. Chen, D. Xiao, Z. Yu, and M. Zhang, "Secure and efficient federated learning via novel multi-party computation and compressed sensing," *Inf. Sci. (Ny).*, vol. 667, p. 120481, 2024.
[23] S. A. Goswami, S. Dave, and K. C. Patel, "Healthcare Informatics Security Issues and Solutions Using Federated Learning," in *Federated Learning for Smart Communication Using IoT Application*, Chapman and Hall/CRC, 2024, pp. 124–154.
[24] M. Suliman and D. Leith, "Two models are better than one: Federated learning is not private for google gboard next word prediction," in *European Symposium on Research in Computer Security*, Springer, 2023, pp. 105–122.
[25] A. Tariq *et al.*, "Trustworthy federated learning: A comprehensive review, architecture, key challenges, and future research prospects," *IEEE Open J. Commun. Soc.*, 2024.
[26] S. R. Abbas, Z. Abbas, A. Zahir, and S. W. Lee, "Federated Learning in Smart Healthcare: A Comprehensive Review on Privacy, Security, and Predictive Analytics with IoT Integration," in *Healthcare*, MDPI, 2024, p. 2587.
[27] C. Zhang, S. Yang, L. Mao, and H. Ning, "Anomaly detection and defense techniques in federated learning: a comprehensive review," *Artif. Intell. Rev.*, vol. 57, no. 6, p. 150, 2024.
[28] N. Bouacida and P. Mohapatra, "Vulnerabilities in federated learning," *IEEe Access*, vol. 9, pp. 63229–63249, 2021.
[29] G. Xia, J. Chen, C. Yu, and J. Ma, "Poisoning attacks in federated learning: A survey," *Ieee Access*, vol. 11, pp. 10708–10722, 2023.
[30] X. Gong, Y. Chen, Q. Wang, and W. Kong, "Backdoor attacks and defenses in federated learning: State-of-the-art, taxonomy, and future directions," *IEEE Wirel. Commun.*, vol. 30, no. 2, pp. 114–121, 2022.
[31] K. N. Kumar, C. K. Mohan, and L. R. Cenkeramaddi, "The impact of adversarial attacks on federated learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 5, pp. 2672–2691, 2023.
[32] B. Liu, N. Lv, Y. Guo, and Y. Li, "Recent advances on federated learning: A systematic survey," *Neurocomputing*, p. 128019, 2024.
[33] S. K. Das, R. Mudi, M. S. Rahman, K. M. Rabie, and X. Li, "Federated Reinforcement Learning for Wireless Networks: Fundamentals, Challenges and Future Research Trends," *IEEE Open J. Veh. Technol.*, 2024.