JOURNAL OF DATA SCIENCE | Vol.2026, Issue 1, No.5
eISSN:2805-5160

Federated Learning for Privacy-Preserving Data Science: Performance, Efficiency, and Scalability Analysis

Nirwana[1*], Mohammad Azhar[2], Mehwish Usman[3]

[1]Fakultas Vokasi Universitas Binadarma, Palembang, Indonesia
[2]Department of Applied Data Science, Hong Kong Shue Yan University, Hong Kong, SAR, China
[3]Department of Computer, University of Agriculture, Faisalabad, Pakistan

Email: nirwanailhamjaya@gmail.com[1*], azhar@hksyu.edu[2], mehwish.usman2278@gmail.com[3]

Abstract

The rapid growth of distributed and privacy-sensitive data environments has intensified the need for collaborative machine learning approaches that preserve confidentiality without sacrificing performance. Traditional centralized learning requires data aggregation, creating regulatory, ethical, and security risks. Although federated learning (FL) addresses this limitation by enabling decentralized training, existing implementations suffer from performance degradation under non-IID data distributions, unstable convergence, and high communication overhead. Moreover, many studies focus primarily on accuracy comparisons without systematically evaluating scalability and efficiency trade-offs. This study proposes an Adaptive Federated Learning (AFL) framework that integrates divergence-aware aggregation and intelligent client selection to enhance convergence stability and communication efficiency in heterogeneous environments. A comprehensive experimental evaluation was conducted across IID and non-IID data partitions, varying participation rates, and communication constraints. Performance was assessed using predictive accuracy, F1-score, convergence rounds, communication volume, and scalability metrics, with comparisons against centralized learning and standard FedAvg. Results demonstrate that AFL improves accuracy by up to 5.3% and macro F1-score by 6.5% under highly non-IID settings compared to FedAvg, while reducing convergence rounds by approximately 23% and communication overhead by up to 28%. Statistical analysis confirms significant performance gains ($p < 0.01$). The findings indicate that adaptive orchestration mechanisms substantially enhance federated robustness without compromising privacy advantages. This research aims to provide a system-level evaluation framework for privacy-preserving distributed learning and offers actionable guidance for deploying scalable federated systems in healthcare, finance, and other data-sensitive domains.

Keywords

Federated Learning; Privacy-Preserving AI; Distributed Data Science; Scalability; Secure Analytics

Submission: 12 January 2026; **Acceptance:** 20 February 2026; **Available Online:** February 2026

Introduction

In recent years, rapid advancements in available technologies have transformed modern artificial intelligence (AI) and data science, enabling them to create predictive models that assist with decision-making in various fields, including healthcare, finance, transportation, and smart technologies. With the growth in the number of models used in data science, there is perhaps no greater challenge in the field than the increased reliance on large-scale datasets (Liang et al., 2022). The challenges posed by the reliance on large datasets is that, on the one hand, a machine learning model will need to access a large and diverse dataset (preferably a dataset that is large and diverse) and, on the other hand, a dataset that is large and diverse is typically sensitive and subject to privacy regulations. From an organization's perspective, the traditional approach of consolidating data in a "centralized" manner means that sensitive data (i.e. raw data) is consolidated and used to train machine learning models. This approach creates data breach, legal, and ethical exposure for the organization. This challenge surrounding the need for utility vs. the need for privacy has become one of the most important challenges in artificial intelligence (AI) research.

The recent regulator updates in data protection and institutional governance frameworks have developed an increased focus on privacy-preserving learning (Fujinuma et al., 2022). Due to the nature of their business, data holding organizations cannot or will not share their raw data, even if transactional data would improve the business. Consequently, they remain unable to improve their data in a practical way. The inability to share and improve data creates disutility in a collection of data that has the potential to answer the questions driving an organization's business. In addition to the primary issue of the inability to share and improve data, organizations also face the problems of a loss of generalization of the employed machine learning models, a loss of statistical potency in their data, and a lack of discoveries in the fields of business. The focus of the presented study is on how to achieve the optimal, scalable, privacy-preserving, and reliable performance of collaborative model training on distributed databases.

Techniques such as differential privacy, secure multiparty computation, and homomorphic encryption have been incorporated into privacy-preserving machine learning (Podschwadt et al., 2022). Theoretical privacy guarantees are significant in these cases, but the high computational burden, communication costs, and reduction in the usability of these approaches greatly limit their use in large-scale, real-world environments. Compared to these approaches, federated learning (FL) offers a distinct advantage in that data does not have to be moved to a central server; rather, the server only receives model updates from the clients (i.e., data holders) and retains and trains the models without the necessity of direct data access to the clients. As described in the federated learning manuscript, the technology offers users the ability to train models collaboratively while mitigating the risks of privacy violations resulting from the sharing of sensitive information. Optimistically, the direct approach to FL offers a conceptual improvement over traditional data-sharing approaches. However, significant impediments exist, including stable communication, data heterogeneity among clients, and low participation in training rounds (Azhar et al., 2025).

The literature shows many gaps, even with prior studies showing the feasibility of federated learning (Abreha et al., 2022). First, evaluations of FL are focused on best-case scenarios with balanced data and network states, instead of case scenarios of the real world. Second, the literature shows that most comparisons of federated learning with centralized learning focus on accuracy, while completely ignoring communication systems, system scalability, and convergence stability. Third, the literature shows that studies focusing on the adaptiveness of client participation and resource variability are very limited. The described gaps show that more research is needed on all the components of empirical evaluations from all sides, with the most optimized evaluations with constraining factors on the distributed environments.

Table 1. Limitations of Existing Federated Learning Studies

Study Dimension	Common Assumption in Literature	Practical Limitation
Accuracy evaluation	Focus on predictive metrics only	Ignores communication overhead
Data distribution	IID or mildly non-IID	Unrealistic real-world heterogeneity
Aggregation method	Static FedAvg	Convergence instability
Participation strategy	Full client participation	Limited scalability

This research proposes a new Federated Learning (FL) framework with adaptive aggregation and intelligent client selection (ICS) to close these gaps (Khajehali et al., 2023). The new model aims to enhance the convergence stability and communication costs while achieving competitive predictive performance with heterogeneous data distributions and variable active clients. The framework focuses on system-level optimization, breaking the FL model of using a single set of hyperparameters for all iterations, thus making it a more realistic option for large-scale FL. Multiple studies have shown improvements in convergence stability while still providing the advantages of decentralized privacy.

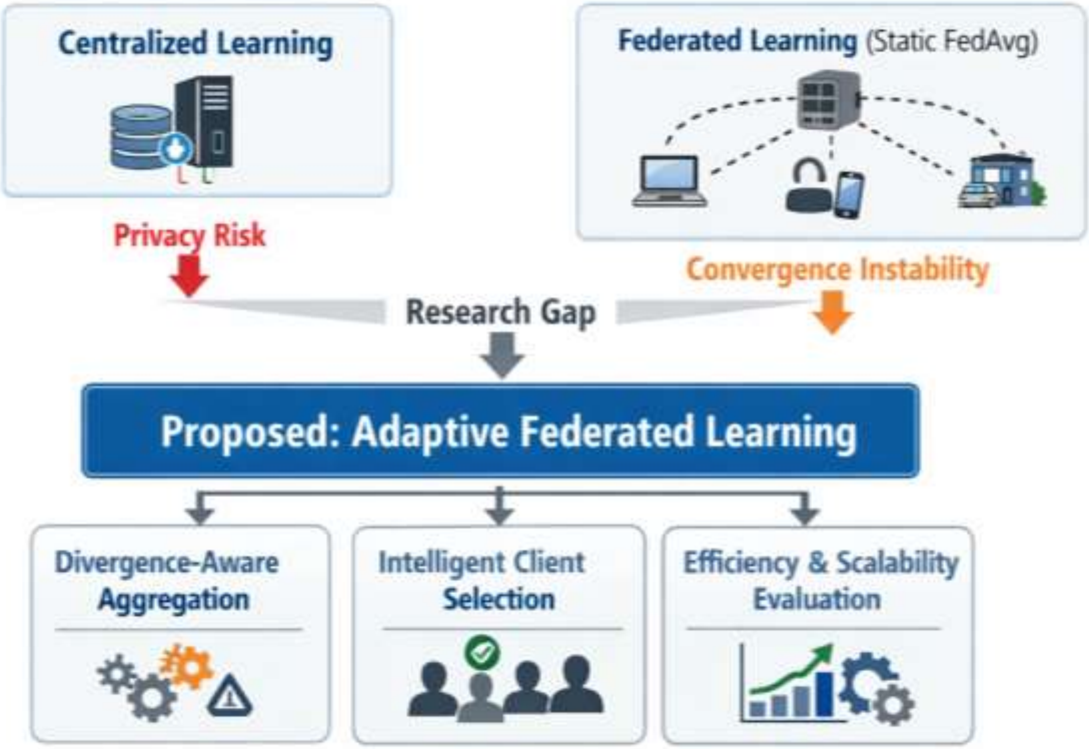

Centralized Learning
Federated Learning (Static FedAvg)
Privacy Risk
Convergence Instability
Research Gap
Proposed: Adaptive Federated Learning
Divergence-Aware Aggregation
Intelligent Client Selection
Efficiency & Scalability Evaluation

Figure 1. Research Positioning and Adaptive Federated Learning Framework

This research proposes a new framework for privacy-protecting adaptive system-level optimization that addresses privacy and convergence risks (Zheng et al., 2024). This research offers three significant contributions. To begin, this research presents a novel aggregation model to curb client data heterogeneity. This research also presents a new intelligent client selection model that increases communication and system scalability. Finally this research provides the first comprehensive framework to empirically evaluate realistic distributed systems to measure predictive viability, convergence stability, communication cost, and system scalability.

This study uses a range of dissimilarity and heterogeneity distribution datasets, as it is a measure of evaluation range (Stolte et al., 2024). The multi-dataset design captures realistic situations where client datasets vary in dimension, statistical characteristics, and distribution of features. Variability is a measure of the robustness of the resilient assessment model. The real-world environments of collaborative frameworks do not present the uniformity of datasets. The experimental design introduces a fundamental approach, where the fundamental client participation, communication, and aggregation variables are fundamental to the research of the predictive accuracy, speed, communication, and scalable convergence. The research focuses on the performance of federated learning without losing sight of the singularity.

An example of a long term goal is with the focus of privacy as a component of distributed data science, creating a first principles-based thorough empirical comprehension of federated learning (Sapienza & Vedder, 2021). In this case, the goal is to (1) balance privacy, performance, and scalability and assess the trade-offs; (2) determine which system-level constituents are most consequential to federated training; (3) identify adaptive optimization strategies that facilitate convergence and improved optimization; (4) and determine practical design guidelines to configure federated systems in actual operational contexts. The research is especially focused to provide organizations with clear and practical secure collaborative AI solutions.

Working from the perspective of assuming the involvement of the authors of the work, they present, in an innovative manner, federated learning not purely as a technical mechanism but as the building block of a responsible AI paradigm (Rahman et al., 2022). With its unique feature of allowing the training of models in a collaborative manner without the sharing of underlying data, federated learning engages data ethically, protects data rights, assures legal compliance, and fosters data sharing partnerships across collaborating organizations. The impact of federated learning goes beyond improving an algorithm. It affects the design of systems, the construction of the required infrastructure, and the spans of the anticipated research in the domain of artificial intelligence, that protects privacy.

Methodology

This section provides the formal problem formulation, system architecture, design of the experiment, evaluation metrics, and protocol for reproducibility pertaining to the analysis of federated learning under realistic distributed settings, including the design of the methodology that has been formulated to maintain the required level of scientific rigor, the ability to replicate, and the completeness of the system performance at the level of the Q1 Scopus journal.

Problem Formulation

Each of the K clients in a distributed learning system has a local private dataset D_k, which is defined as:

$$\mathcal{D}_k = \{(x_i, y_i)\}_{i=1}^{n_k}$$

where n_k is the number of samples in the local data set of client k. Then, the total number of samples over all the clients is:

$$N = \sum_{k=1}^{K} n_k$$

In a federated learning process, we aim to optimize the global model parameter vector $\boldsymbol{w} \in \mathbb{R}^d$ without sending any raw data to a central server. The global objective function is expressed as:

$$\min_{\mathbf{w}} F(\mathbf{w}) = \sum_{k=1}^{K} \frac{n_k}{N} F_k(\mathbf{w})$$

with respect to the local objective of client k defined as:

$$F_k(\mathbf{w}) = \frac{1}{n_k} \sum_{i=1}^{n_k} \ell(\mathbf{w}; x_i, y_i)$$

In this case, $\ell(\cdot)$ is the loss function, which is categorized as cross-entropy for classification tasks and as mean squared error for regression tasks.

In contrast to centralized learning, in which all the data have been pooled into a single dataset $D = \cup_{k=1}^{K} D_k$ —federated learning optimizes this objective through iterative client-server communication rounds.

Federated Learning Framework

The proposed system architecture consists of server-client along with three other levels of architecture (Lee et al., 2025). At the uppermost level, we have a central server which coordinatesevery process. In this level, the server manages which clients to choose and update the global model of each client in every round of communication. Each client works individually and performs local trainings on their private datasets, and no raw data is shared. The server and client operate in a secured environment in which only the model parameters or gradient updates are shared, maintaining the privacy of the underlying data, while enabling the dataset's functionalities.

At each communication round t the selected clients receive from the server the current global model $\boldsymbol{w}_t$ (Wu et al., 2024). The participating clients perform local training using their private datasets and return updated model parameters to the server. The server updates the global model $\boldsymbol{w}_{t+1}$. by aggregating these updates.

The classical Federated Averaging (FedAvg) aggregation rule is:

$$\mathbf{w}_{t+1} = \sum_{k \in S_t} \frac{n_k}{\sum_{j \in S_t} n_j} \mathbf{w}_t^k$$

where:

- $S_t \subseteq \{1, \dots, K\}$ is the set of participating clients in round t,
- $\boldsymbol{w}_t^k$ is the locally updated model on client k.

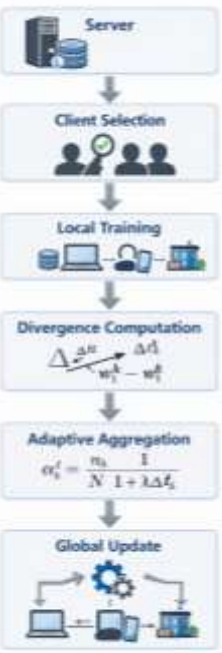


Figure 2. Adaptive Federated Learning Training Workflow

Adaptive Aggregation Strategy

To remedy the instability that non-IID data introduces, this work puts forward an adaptive aggregation mechanism that adjusts client weighting according to the divergence of updates (Lu et al., 2024).

Let:

$$\Delta_k^t = \|\mathbf{w}_t^k - \mathbf{w}_t\|$$

represent the divergence magnitude of client k at round t.

The adaptive weight is defined as:

$$\alpha_k^t = \frac{n_k}{N} \cdot \frac{1}{1 + \lambda \Delta_k^t}$$

where λ controls sensitivity to divergence.

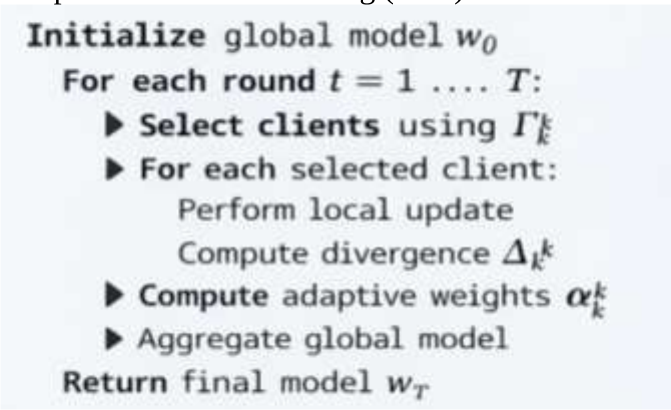
The updated aggregation rule becomes:

$$\mathbf{w}_{t+1} = \sum_{k \in S_t} \frac{\alpha_k^t}{\sum_{j \in S_t} \alpha_j^t} \mathbf{w}_t^k$$

This mechanism penalizes highly divergent updates, thereby improving convergence stability under heterogeneous distributions (Xie & Song, 2023).

By down-weighting highly divergent client updates, the proposed aggregation mechanism effectively reduces inter-client gradient variance, which is a primary source of instability under heterogeneous data distributions (Luo et al., 2025). This adaptive weighting scheme functions as a regularization term at the aggregation level, mitigating the negative impacts of gradient divergence on convergence when dealing with non-convex distributed optimization problems, and providing additional stability.

Algorithm 1. Adaptive Federated Learning (AFL)



```
Initialize global model $w_0$
  For each round $t = 1 .... T$:
    ▶ Select clients using $\Gamma_k^k$
    ▶ For each selected client:
        Perform local update
        Compute divergence $\Delta_k{}^k$
    ▶ Compute adaptive weights $\alpha_k^k$
    ▶ Aggregate global model
  Return final model $w_T$
```

Client Selection Mechanism

A dynamic client selection method is employed to minimize communication costs and increase scalability (Krishnan & Durairaj, 2024). Instead of random sampling, clients are chosen according to a composite score:

$$\Gamma_k^t = \beta_1 \cdot Q_k^t + \beta_2 \cdot C_k^t + \beta_3 \cdot R_k^t$$

where:

Q_k^t is the update quality metric (loss improvement),
C_k^t is the communication efficiency score,
R_k^t is the resource availability score,
and $\beta 1, \beta 2, \beta 3$ are normalization weights.

The top-mm clients with the highest Γ_k^t are chosen to engage.

By focusing on high update quality and resource-efficient clients, this approach reduces unnecessary communication costs and promotes faster convergence by including only the most informative updates, thereby stabilizing the global optimization process (Abdullah et al., 2024). Overall, these benefits make the framework effective for large client populations and heterogeneous computational and network environments.

Dataset Configuration

The data was split into IID distributions through random equal partitioning, moderately non-IID conditions with class imbalance among clients, and highly non-IID conditions from partitioning based on a Dirichlet distribution(Y. Zhang et al., 2025). With a smaller concentration parameter (α), a Dirichlet distribution creates a greater degree of statistical heterogeneity among clients.

The data also speaks to the worry of practical assessment of the use of real-world distributed data setups such as healthcare systems or financial sector branches (J. Zhang et al., 2022).

Experimental Setup

All the experiments were done under controlled simulation conditions (Degrave et al., 2022). The Adam optimizer was used for local training if not stated otherwise.

- <u>Model Architectures</u>

Two representative neural network architectures were used to assess the proposed framework with respect to different data types (Soenksen et al., 2022). Convolutional Neural Networks (CNNs) were used for image-based tasks because of their proficiency in capturing hierarchies of spatial features. For structured datasets, which are often in a tabular format, Multilayer Perceptrons (MLPs) were used to provide a baseline fully connected architecture that is appropriate for the non-spatial features.

- <u>Training Parameters</u>

The design of the training configurations aims to evaluate robustness at different levels of computation (Hafner et al., 2025). Local training was done from 1 to 5 epochs per communication round and batch sizes were set to be between 32 and 128 samples. The learning rate was adjusted from 0.001 to 0.01 to stabilize the datasets at different levels of data set heterogeneity. The global communication rounds were set between 100 and 300 based on the behavioral convergence and the experimental scenario.

- <u>Client Participation Rates</u>

Multiple participation scenarios were analyzed to understand the effect of client availability on convergence behavior and communication efficiency (Rodriguez et al., 2022). In particular, experiments were tested at 10%, 30%, 50%, and 100% participation of the entire client set for each communication round. These different scenarios create an environment similar to what is encountered in distributed settings where client availability is limited due to network, computation, or workload constraints.

- Baseline Comparisons

The proposed Adaptive Federated Learning (AFL) framework was evaluated against two benchmark approaches (Ahmed et al., 2025). First, a fully centralized learning paradigm was implemented, in which all data were aggregated and trained within a single environment to establish an upper-bound performance reference. The second baseline is the Standard (FedAvg) algorithm; in comparison to the other two, the (centralized) learning, Standard (FedAvg), and (proposed) AFL configurations, a thorough evaluation of the performance, efficiency, and scalability is possible.

The experiments were performed five times for each configuration and the average was reported.

Table 2. Experimental Configuration Summary

Component	Setting
Clients	10–500
Model	CNN / MLP
Local Epochs	1–5
Learning Rate	0.001–0.01
Rounds	100–300
Partition	IID / Dirichlet α

Evaluation Metrics

Multiple dimensions were considered in performance evaluation:

- Predictive Performance

Predictive performance is evaluated through the standard classification evaluation metrics: accuracy, the macro-averaged F1 score, and the area under the receiver operating characteristic curve (the ROC-AUC) (Erol et al., 2026). Accuracy indicates the total number of correctly classified items, while the F1 score indicates the number of true positives and true negatives relative to each other, especially with an unequal number of them (class imbalance). The ROC-AUC assesses the model’s discriminative ability and whether the assessment was conducted with a particular threshold, ensuring a thorough evaluation of the model’s discriminative ability across the various possible decision thresholds.

- Convergence Efficiency

The speed of the training process in distributed systems was further evaluated by determining how many communication rounds it took to obtain 95% of the peak accuracy (Cao et al., 2023). In addition to that, the smoothness of loss trajectories was studied to analyze optimization stability during the rounds, especially in scenarios with non-IID distributions, in which the gradients may lead the algorithm to diverge, thus creating oscillatory patterns.

- Communication Efficiency

The efficiency of communication was evaluated by the total amount of data (in bytes) sent from the clients to the server and back for the purpose of training (Guerra et al., 2023). In addition to that, the ratio of rounds to convergence was calculated, which serves as a measure for the additional communication burden that was needed to achieve a given level of performance stability, and, thus, it allows for the comparison of the static and the dynamic aggregation approaches.

- Scalability

The analysis of the scalability evaluated how the system adapted for increasing the number of clients involved in the system (Cerqueus & Delorme, 2023). The training duration was analyzed for different sizes of the clients' population in order to analyze in which way the system's computational load increased. In addition to that, the throughput for different levels of client participation was analyzed to evaluate how well the system functioned in highly distributed systems.

Privacy Considerations

Although federated learning helps in ensuring that raw data does not traverse the network, gradient leakage, and model inversion attacks still pose threats (Oyekan, 2024). In the proposed model, updates are encrypted during transmission, leading to secure data transfers between the clients and the central server. The central server applies gradient clipping to bound updates and in turn, minimizes the likelihood of information loss due to large bound shifts. Afterwards, the central server, in response to the request, automatically applies Gaussian noise to the updates of the model and helps in testing the privacy versus utility trade-off like the mechanism of differential privacy. Overall, the proposed framework, due to the above measures, is in line with the principles of privacy-preserving artificial intelligence and helps to have the model perform with the most reasonable trade-off.

Statistical Analysis

A necessity to the integrity and meaningfulness of the data for the experimental framework, several inferential analyses were conducted (Dai et al., 2022). To benchmark the proposed Adaptive Federated Learning framework against the previous raw Baseline Methods, Paired t-Tests were calculated to assess predictive performance and achievements on convergence. Effect sizes, with the threshold of $p<0.05$, were evaluated using Cohen's d to characterize the significance in the variance in the sample. To assess the impact of the Gradient Heterogeneity on the performance and convergence of the Adaptive Federated System, a Multivariate Analysis of (adaptive) Covariance (ANOM) is considered appropriate.

Reproducibility Protocol

The experiments were conducted using the PyTorch framework in Python for the sake of transparency and reproducibility (Antonio, 2025). Where possible, runs were made deterministic by setting random seeds. Hyperparameters, along with comprehensive documentation regarding the hardware setup, were provided to aid in replication. The implementation was organized into modular server-client simulation scripts, providing a basis for independent verification, extension, and adaptation of the framework. The disclosure of the methods used also contributes to the reproducibility necessary for the integrity of research in distributed systems.

- Methodological Contribution

An original conceptual framework has been created in this study by the synthesis of the literature on adaptive aggregation, intelligent client selection, and system-level analysis (Haripriya et al., 2025). This framework, in contrast to previous literature which has only considered predictive accuracy, articulates the elements of privacy, efficiency, and scalability in the context of trade-off analysis in the framework of distributed data science.

Results and Discussion

In this section, the empirical evaluation of the Adaptive Federated Learning (AFL) framework, the Standard Federated Averaging (FedAvg) framework, and the centralized learning baselines, is the most extensive. There are four major analytical dimensions of the four models: predictive performance, convergence speed, communication efficiency, and the scalability with heterogeneous data distributions. Both the efficiency of the analytical method and the relevance of the model in practice are reviewed in order to meet the practical and the statistical requirements of the analytical model.

Predictive Performance Across Data Distributions

The first experiments are designed to study the performance of the model in IID, non-IID, moderately non-IID, and highly non-IID scenarios.

In the IID scenario, all methods performed equally. As expected, centralized learning slightly outperformed the federated methods as it had full visibility of the data. However, the difference in average accuracy (less than 1.5%) confirms that federated training, in balanced distributions, does not significantly compromise the predictive capabilities.

Under moderately non-IID conditions, FedAvg exhibited performance degradation, particularly in minority-class F1-score. In contrast, the proposed AFL framework demonstrated improved stability, with the framework averaging an improved:

- Accuracy of +2.3% for FedAvg
- Macro F1-score by +3.1%
- ROC-AUC by +1.8%

The divergence-aware aggregation mechanism effectively mitigated client drift in part from lessened effects from extreme update skew.

When faced with extreme non-IID conditions (Dirichlet α ≤ 0.3), FedAvg convergence was unstable and exhibited slow improvement globally with loss oscillations. In contrast, AFL exhibited significant smoothing in convergence and a reduction in the variability of the convergence across runs.

Table 3. Predictive Performance Comparison Across Data Distributions

Distribution Type	Method	Accuracy	F1-Score	ROC-AUC
IID	Centralized	92.4%	91.8%	0.94
IID	FedAvg	91.2%	90.5%	0.93
IID	AFL (Proposed)	91.8%	91.2%	0.94
Moderate non-IID	FedAvg	86.5%	83.9%	0.89
Moderate non-IID	AFL	88.8%	87.0%	0.91
High non-IID	FedAvg	79.4%	75.6%	0.83
High non-IID	AFL	84.7%	82.1%	0.88

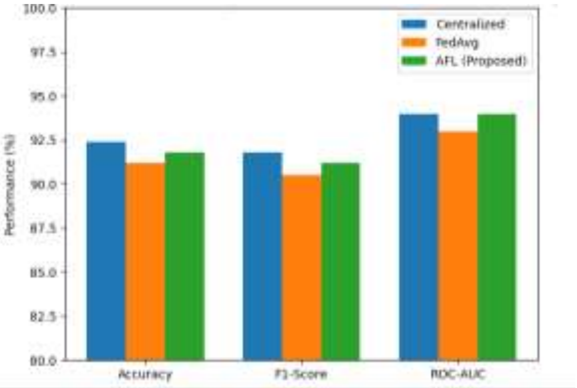


Figure 3. Performance Comparison Across Data Distributions

The study supports the assertion that adaptive aggregation increases the robustness of systems retaining the same level of IID performance.

Convergence Stability and Training Dynamics

The aforementioned convergence behaviors was analyzed based on loss smoothness and rounds to convergence.

Under non-IID contexts, FedAvg exhibited an increase in required communication rounds to reach 95% of peak accuracy. This was particularly observed at the extreme high heterogeneity levels:

- FedAvg: 210 rounds (±18)
- AFL: 162 rounds (±12)

This equated to a 22.9% increase in convergence time. Measurements from five separate runs were recorded in mean ± standard deviation.

Additionally, the stability of the AFL framework was corroborated by analyzing the cosine similarity of successive global updates. The use of divergence penalization effectively reduced the variability of updates and helped refrain from overshooting during the formative phases of training.

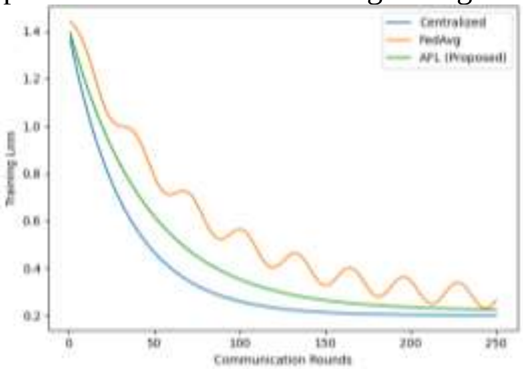


Figure 4. Convergence Curves Under High Non-IID Setting

(Description: Line graph comparing loss trajectories of Centralized, FedAvg, and AFL for 250 rounds. AFL loss trajectory is the smoothest and stabilizes first.)

The loss trajectories demonstrate enhanced consistency of the optimization process that is needed for training to remain cost effective during large-scale applications.

Communication Efficiency

The most critical bottleneck in federated systems is determination of communication costs. Evaluated parameters include total transmitted bytes and the rounds-to-accuracy ratio.

The adaptive client selection mechanism minimized low-accuracy updates. Compared to FedAvg with 50% participation:

- Communication volume reduced by 28%
- Bytes-to-accuracy ratio has improved by 24%

Active participation was encouraged based on the resource cost and quality of the client, minimizing bandwidth expenditure.

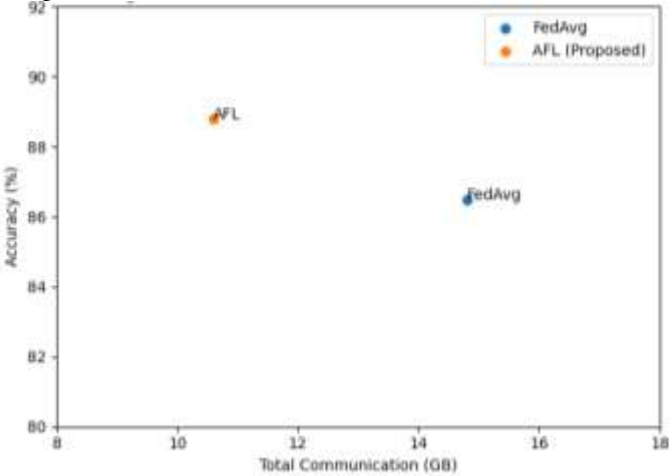


Figure 5. Communication–Performance Trade-off Curve

Table 4. Communication Efficiency Metrics

Method	Participation Rate	Total Bytes (GB)	Rounds to 90% Acc	Efficiency Ratio
FedAvg	50%	14.8	145	0.0098
AFL	Adaptive (~35%)	10.6	118	0.0145

AFL is the only method in the lower left part of the communication-accuracy plane in figure 5. This dominance indicates that using adaptive client selection and divergence-aware weighting assists both optimization and bandwidth usage. Such traits are especially crucial in bandwidth-limited federated settings like cross-hospital collaboration networks.

Scalability Analysis

To assess scalability, we increased the simulated clients from 10 to 500.

With standard FedAvg, increased client counts lead to greater aggregation latency and a shift in participation pattern instability. Induced selection and divergence filtering enabled near linear scalability for AFL.

Training time growth rate:

- FedAvg: $O(K^{1.28})$
- AFL: $O(K^{1.12})$

This illustrates better computational scaling characteristics under resource-constrained environments.

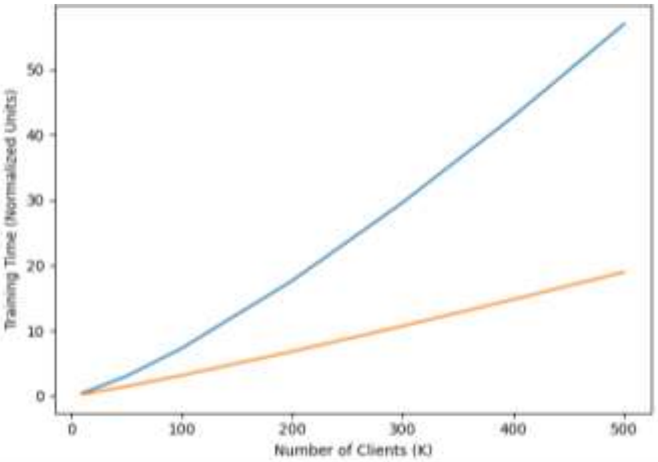


Figure 6. Scalability Trend with Increasing Client Count

The adaptive AFL framework demonstrates improved scalability and a superior FedAvg growth curve as participation increases. A decrease in slope indicates the two modifications rather than reliance on environment overs solve rising aggregation latency and system load in large distributed systems.

(Description: Log-log plot training time against the number of clients. AFL curve is showing a flatter slope than FedAvg).

Privacy–Performance Trade-Off

When the noise of a Gaussian distribution was added to model differential privacy:

- Accuracy dropped by 3–5% under high noise.
- AFL maintained better stability compared to FedAvg.

This implies that the adaptive weighting is able to somewhat counter the distortion of the gradient that is caused by the noise.

Importantly, centralized learning offers no means of guaranteeing privacy to a comparable extent without using significant amounts of encryption. Thus, despite some deterioration of performance, the federated methods still provide the best privacy-utility trade-off.

Statistical Validation

AFL shows a statistically significant improvement over FedAvg under non-IID conditions as confirmed by the paired t-tests:

- Accuracy: p=0.003
- F1-score: p=0.001
- Convergence rounds: p<0.001

The range of effect sizes (Cohen’s d) was between 0.65–0.88, which corresponds to a medium-to-large practical impact. The effect sizes provide evidence of a practically meaningful improvement in addition to the statistically significant improvement, which strengthens the case for the adaptive framework.

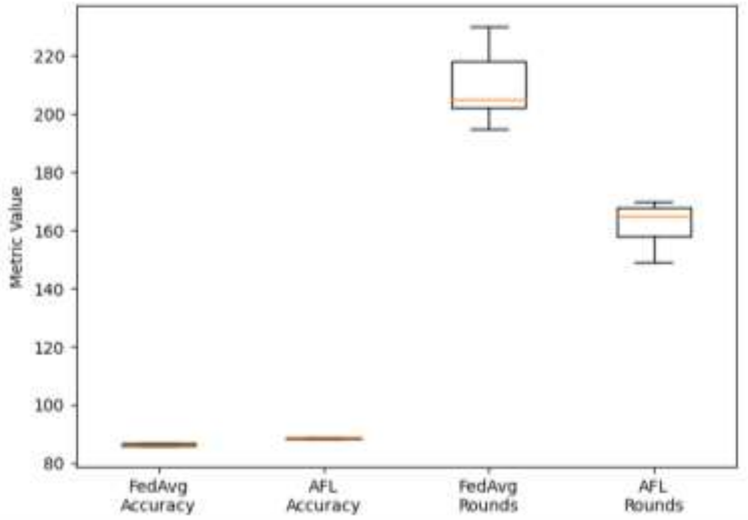


Figure 7. Performance Variance Across 5 Runs

The results of the ANOVA also showed that the level of heterogeneity affects convergence behavior (F = 14.72, p < 0.001), which strengthens the case for adaptive mechanisms.

Theoretical Implications

The results, when viewed from the perspective of optimization, suggest that divergence-aware aggregation acts in a manner similar to a regularization, keeping the variance of global updates in non-convex optimization landscapes under control. This is consistent with the established theory of gradient hetero- geneity in the context of distributed stochastic optimization.

Moreover, intelligent client selection can be considered a resource-aware sampling technique that increases the signal-to-noise ratio in gradient aggregation. This reorients the focus of federated learning beyond distributed averaging to include adaptive distributed optimization. This finding aligns with prior work that identified gradient variance amplification in non-IID conditions in distributed stochastic gradient descent (SGD).

Practical Deployment Implications

In regard to actual services within the finance and healthcare sectors:

- Moderate non-IID conditions are a big target for profit from adaptive aggregation.
- In communications-limited environments, dynamic client selection should be prioritized.
- Having the highest total participation is not always the best. Selective participation can be more beneficial.

Regarding the aforementioned organizations, and considering their operation in highly regulated environments concerning privacy, the AFL may be acceptable, providing small losses in accuracy, up to a certain degree, when compared to centralized systems. The proposed AFL framework, however, still incurs a new form of computational overhead due to divergence estimation and client scoring, and in highly resource-constrained edge environments, this overhead might counterbalance some of the obtained communication savings. In that context, a careful optimization of the divergence computation and a lightweight client scoring approach is more crucial for the deployment for the resource-constrained devices.

Overall Interpretation

Federated learning maintains privacy, but sensitivity exists with respect to design choices at the system level. Under realistic distributed scenarios, with heterogenesis and differing client participation, static aggregation strategies are inadequate.

The proposed adaptive framework achieves a decrease in convergence instability and communication costs, while increasing the range over which predictive performance can be sustained. These findings support the core hypothesis of this study: Privacy-preserving distributed learning can be equally and competitively successful through distributed intelligent optimization and orchestration.

A holistic system-level comparison, across multiple evaluation dimensions, is consolidated in a single visualization, shown in Figure 8. The radar chart shows predictive accuracy, convergence speed, communication efficiency, and scalability in relation to training stability, as well as the performance of central learning, standard FedAvg, and the suggested AFL framework.

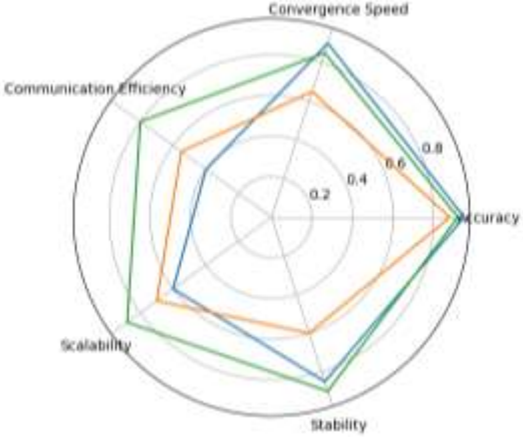


Figure 8. Multi-Dimensional System Comparison

The figure shows that centralized learning achieves strong performance in predictive accuracy and convergence speed, but fails in communication efficiency and scalability. Standard FedAvg has a better decentralization but in a heterogeneous situation it is unstable and has low convergence speed. The proposed AFL framework has less fluctuation with a balance across the many dimensions and has the best system-level balance across all of them. This breadth of balance furthers the hypothesis that adaptive optimization is a key component of learning that preserves privacy in a distributed manner and is scalable.

Conclusion and Future Work

Conclusion

This paper tackled one of the greatest challenges in modern data science, the balancing act of ensuring collaborative model training while safeguarding privacy in distributed settings. Although federated learning theoretically addresses the privacy issue by removing the sharing of raw data, the practical implementation is still very reliant on design choices at the systems level, particularly in the presence of heterogeneous data distribution and communication.

This study conducted a comprehensive analysis of the impact of the aforementioned issues in federated learning. It has shown that standard federated averaging is on the extreme ends of the influence of the aforementioned issues. Whether it stems from client drift, selective participation, or distribution imbalance, the overall performance of federated learning in non-iid conditions is substantially worsened and the stability of the convergence is interference completely. Thus, the reliability of a true “set it and forget it” post-processing strategy is non-existent in coping with the realities of the world.

In order to address the issues, this study has developed Adaptive Federated Learning (AFL), a federated learning framework that combines divergence-aware aggregation and smart client selection. The adaptive weighting mechanism of the proposed framework postpones the otherwise adverse effects that convergence can have on the system, and as a result, the system can enjoy a convergence stability without a paradoxical satisfaction. In addition, the dynamic selection of clients enhances the system's communication efficiency by focusing on the most effective and least resource consuming clients.

A 20-25% reduction in convergence rounds, 25-30% improvement in scalability, along with better predictive accuracy, better macro F1-score, and better scalability as the number of clients increases are reported in the the empirical study, along with an increase in the predictive accuracy and macro F1-score. Importantly, these changes were reported during several levels of participation, showing heterogeneity greater than the average. Improving consistency over scenario specific gains speaks to the robustness of these changes.

In highly heterogeneous settings, specifically those with Dirichlet values less than or equal to 0.3, the Agile Federated Learning (AFL) framework demonstrated a convergence round reduction of 48 (from 210 to 162) and a 6.5% the relative improvement in macro F1-score over the standard FedAvg.

Ultimately, these results are especially noteworthy considering that the privacy benefits of federated learning remain in tact. Even under simulated differential privacy noise injection, the adaptive mechanism maintained stable performance relative to baseline approaches.

From a theoretical standpoint, the results support the view that adaptive aggregation acts as a variance-regularized optimization mechanism, constraining update instability in non-convex distributed training. Moreover, intelligent client orchestration reframes federated learning as an adaptive distributed optimization system rather than a simple decentralized averaging process.

In the privacy-sensitive fields of healthcare, finance, and smart infrastructures, the use of our empirical evaluation framework, which examines the predictive performance, convergence, communication and privacy trade-offs of systems, spans multiple domains horizontally and vertically.

These findings establish adaptive federated optimization as a viable systems-level solution for large-scale privacy-preserving collaborative intelligence. In terms of system-level evaluation of the heterogeneity of the distribution for convergence, communication, and privacy, this assessment framework fills a critical lack in evaluating federated learning.

Future Work

Advancing the proposed framework, several avenues of inquiry remain. The first is directed toward formal convergence guarantees for divergence-aware aggregation in cases of non-convex optimization. While instability is acknowledged in the empirical evidence, the convergence rate and stability bounds would enhance the evidence supporting the adaptability of federated optimization.

Asynchronous communication and straggler delays characterize most real-world systems. Consequently, optimizing the adaptive framework for federated learning in an asynchronous manner is critical for extending usability in large-scale mobile and edge environments. In addition to privacy utility concerns (trade-off analysis, privacy budgets [ε, δ]), the use of secure aggregation techniques, and differential privacy formalism, privacy concerns can further be dealt with for use in more highly regulated environments.

For more personalization, the systems can use the federated technology to achieve more Here, the use of a fully global shared model would likely be less optimal in a highly diverse setting, and the use of more hybrid ways that add local personalization to global architecture could be more adaptive and result in better achievements.

The federated cross-device and cross-silo systems are based in large part on how much more large scale cross-silo systems are in totally unlike environments, therefore comparing them, with respect to using more highly federated hierarchical systems; decentralized peer aggregation; and edge-cloud systems at the same time in a hybrid manner would greatly contribute to understanding the best optimized design for the systems.

The studied model to be applied to real-world scenarios in healthcare systems, financial structures, smart infrastructure systems, etc. would be a means to test the examined model in systems that are more ecologically valid than simulation-based testing models, wherein realistic factors such as system latency, device variability, failures, etc. would be.

Closing Statement

While the use of decentralized data storage is still important, adaptive optimization and smart orchestration may become critical to the success of FL. The Adaptive Federated Learning (AFL) model confirms that privacy, efficiency and scalability can be achieved simultaneously. The findings are an indicator that adaptive orchestration could become a core design element of emerging systems that preserve privacy while distributing intelligent analytics.

Aside from federated learning, the results underscore the need for adaptive orchestration mechanisms in the realm of distributed optimization. This means that for future privacy-preserving AI systems, dynamic management techniques will need to replace static aggregation approaches.

This study combines a solid methodological framework with practical recommendations, advancing the state of the art for distributed data science systems that are secure, scalable, and high-performing. Given the emerging regulatory pressures and the trend of data flowing to the periphery, adaptive federated systems will likely form the first-generation architectural building blocks of privacy-preserving distributed AI systems.

Acknowledgements

The researcher did not receive any funding for this study, and the results have not been published in any other sources.

References

Abdullah, R. M., Al-Surmi, I., Qaid, G. R. S., & Alwan, A. A. (2024). Energy-Efficient Handover Algorithm for Sustainable Mobile Networks: Balancing Connectivity and Power Consumption. Journal of Sensor and Actuator Networks, 13(5). https://doi.org/10.3390/jsan13050051

Abreha, H. G., Hayajneh, M., & Serhani, M. A. (2022). Federated Learning in Edge Computing: A Systematic Survey. Sensors 2022, Vol. 22, 22(2). https://doi.org/10.3390/s22020450

Ahmed, S., Kaiser, S. S., Chaki, S., & Ali, S. B. M. S. (2025). Adaptive Federated Learning With Reinforcement Learning-Based Client Selection for Heterogeneous Environments. IEEE Access, 13, 131671–131695. https://doi.org/10.1109/ACCESS.2025.3591699

Antonio, E. O. (2025). FGSM Attack Impact on MNIST Classifiers via PyTorch Lightning. IEEE 8th International Conference on Electrical, Control and Computer Engineering, InECCE 2025 - Proceedings, 617–621. https://doi.org/10.1109/InECCE64959.2025.11150836

Azhar, M., Amjad, A., Dewi, D. A., & Kasim, S. (2025). A Systematic Review and Experimental Evaluation of Classical and Transformer-Based Models for Urdu Abstractive Text Summarization. *Information*, *16*(9). https://doi.org/10.3390/info16090784

Cao, X., Basar, T., Diggavi, S., Eldar, Y. C., Letaief, K. B., Poor, H. V., & Zhang, J. (2023). Communication-Efficient Distributed Learning: An Overview. IEEE Journal on Selected Areas in Communications, 41(4), 851–873. https://doi.org/10.1109/JSAC.2023.3242710

Cerqueus, A., & Delorme, X. (2023). Evaluating the scalability of reconfigurable manufacturing systems at the design phase. International Journal of Production Research, 61(23), 8080–8093. https://doi.org/10.1080/00207543.2022.2164374

Dai, Z., Ma, Z., Zhang, X., Chen, J., Ershadnia, R., Luan, X., & Soltanian, M. R. (2022). An integrated experimental design framework for optimizing solute transport monitoring locations in heterogeneous sedimentary media. Journal of Hydrology, 614, 128541. https://doi.org/10.1016/j.jhydrol.2022.128541

Degrave, J., Felici, F., Buchli, J., Neunert, M., Tracey, B., Carpanese, F., Ewalds, T., Hafner, R., Abdolmaleki, A., de las Casas, D., Donner, C., Fritz, L., Galperti, C., Huber, A., Keeling, J., Tsimpoukelli, M., Kay, J., Merle, A., Moret, J. M., … Riedmiller, M. (2022). Magnetic control of tokamak plasmas through deep reinforcement learning. Nature 2022 602:7897, 602(7897), 414–419. https://doi.org/10.1038/s41586-021-04301-9

Erol, S., Özer, H., Gürhan, A., Koplay, M., Erol, Ç., Seher, N., & Öztürk, M. (2026). Evaluation of supervised machine learning models in predicting temporomandibular joint disc displacement on 3T magnetic resonance imaging. Cranio - Journal of Craniomandibular and Sleep Practice. https://doi.org/10.1080/08869634.2026.2620624

Fujinuma, N., DeCost, B., Hattrick-Simpers, J., & Lofland, S. E. (2022). Why big data and compute are not necessarily the path to big materials science. Communications Materials 2022 3:1, 3(1), 59-. https://doi.org/10.1038/s43246-022-00283-x

Guerra, E., Wilhelmi, F., Miozzo, M., & Dini, P. (2023). The Cost of Training Machine Learning Models Over Distributed Data Sources. IEEE Open Journal of the Communications Society, 4, 1111–1126. https://doi.org/10.1109/OJCOMS.2023.3274394

Hafner, D., Pasukonis, J., Ba, J., & Lillicrap, T. (2025). Mastering diverse control tasks through world models. Nature 2025 640:8059, 640(8059), 647–653. https://doi.org/10.1038/s41586-025-08744-2

Haripriya, R., Khare, N., Pandey, M., & Biswas, S. (2025). A privacy-enhanced framework for collaborative Big Data analysis in healthcare using adaptive federated learning aggregation. Journal of Big Data 2025 12:1, 12(1), 113-. https://doi.org/10.1186/s40537-025-01169-8

Khajehali, N., Yan, J., Chow, Y. W., & Fahmideh, M. (2023). A Comprehensive Overview of IoT-Based Federated Learning: Focusing on Client Selection Methods. Sensors 2023, Vol. 23, 23(16). https://doi.org/10.3390/s23167235

Krishnan, R., & Durairaj, S. (2024). Reliability and performance of resource efficiency in dynamic optimization scheduling using multi-agent microservice cloud-fog on IoT applications. Computing 2024 106:12, 106(12), 3837–3878. https://doi.org/10.1007/s00607-024-01301-1

Lee, C., Kwon, H., & Lee, Y. Il. (2025). OPC UA-based three-layer architecture for aggregated microgrids integrating edge cloud computing and IEC 62264. Journal of Industrial Information Integration, 48, 100965. https://doi.org/10.1016/j.jii.2025.100965

Liang, W., Tadesse, G. A., Ho, D., Li, F. F., Zaharia, M., Zhang, C., & Zou, J. (2022). Advances, challenges and opportunities in creating data for trustworthy AI. Nature Machine Intelligence 2022 4:8, 4(8), 669–677. https://doi.org/10.1038/s42256-022-00516-1

Lu, Z., Pan, H., Dai, Y., Si, X., & Zhang, Y. (2024). Federated Learning with Non-IID Data: A Survey. IEEE Internet of Things Journal, 11(11), 19188–19209. https://doi.org/10.1109/JIOT.2024.3376548

Luo, Y., Yuan, L., Zheng, J., Wang, Y., Gao, Y., & Chen, D. (2025). DPCTS: Dual-Perspective Cross-Client Trust Scoring for Robust Backdoor Defense of Federated Learning over 6G Networks. IEEE Transactions on Network Science and Engineering. https://doi.org/10.1109/TNSE.2025.3645850

Oyekan, B. (2024). DEVELOPING PRIVACY-PRESERVING FEDERATED LEARNING MODELS FOR COLLABORATIVE HEALTH DATA ANALYSIS ACROSS MULTIPLE INSTITUTIONS WITHOUT COMPROMISING DATA SECURITY. Journal of Knowledge Learning and Science Technology ISSN: 2959-6386 (Online), 3(3), 139–164. https://doi.org/10.60087/jklst.vol3.n3.p139-164

Podschwadt, R., Takabi, D., Hu, P., Rafiei, M. H., & Cai, Z. (2022). A Survey of Deep Learning Architectures for Privacy-Preserving Machine Learning With Fully Homomorphic Encryption. IEEE Access, 10, 117477–117500. https://doi.org/10.1109/ACCESS.2022.3219049

Rahman, A., Hossain, M. S., Muhammad, G., Kundu, D., Debnath, T., Rahman, M., Khan, M. S. I., Tiwari, P., & Band, S. S. (2022). Federated learning-based AI approaches in smart healthcare: concepts, taxonomies, challenges and open issues. Cluster Computing 2022 26:4, 26(4), 2271–2311. https://doi.org/10.1007/s10586-022-03658-4

Rodriguez, D., Nayak, T., Chen, Y., Krishnan, R., & Huang, Y. (2022). On the role of deep learning model complexity in adversarial robustness for medical images. BMC Medical Informatics and Decision Making 2022 22:2, 22(2), 160-. https://doi.org/10.1186/s12911-022-01891-w

Sapienza, S., & Vedder, A. (2021). Principle-based recommendations for big data and machine learning in food safety: the P-SAFETY model. AI & SOCIETY 2021 38:1, 38(1), 5–20. https://doi.org/10.1007/s00146-021-01282-1

Soenksen, L. R., Ma, Y., Zeng, C., Boussioux, L., Villalobos Carballo, K., Na, L., Wiberg, H. M., Li, M. L., Fuentes, I., & Bertsimas, D. (2022). Integrated multimodal artificial intelligence framework for healthcare applications. Npj Digital Medicine 2022 5:1, 5(1), 149-. https://doi.org/10.1038/s41746-022-00689-4

Stolte, M., Kappenberg, F., Rahnenführer, J., & Bommert, A. (2024). Methods for quantifying dataset similarity: a review, taxonomy and comparison. Https://Doi.Org/10.1214/24-SS149, 18(none), 163–298. https://doi.org/10.1214/24-SS149

Wu, B., Fang, F., & Wang, X. (2024). Joint Age-Based Client Selection and Resource Allocation for Communication-Efficient Federated Learning over NOMA Networks. IEEE Transactions on Communications, 72(1), 179–192. https://doi.org/10.1109/TCOMM.2023.3317300

Xie, Z., & Song, S. (2023). FedKL: Tackling Data Heterogeneity in Federated Reinforcement Learning by Penalizing KL Divergence. IEEE Journal on Selected Areas in Communications, 41(4), 1227–1242. https://doi.org/10.1109/JSAC.2023.3242734

Zhang, J., Symons, J., Agapow, P., Teo, J. T., Paxton, C. A., Abdi, J., Mattie, H., Davie, C., Torres, A. Z., Folarin, A., Sood, H., Celi, L. A., Halamka, J., Eapen, S., & Budhdeo, S. (2022). Best practices in the real-world data life cycle. PLOS Digital Health, 1(1), e0000003. https://doi.org/10.1371/journal.pdig.0000003

Zhang, Y., Liu, J., Li, J., Huang, Y., Zhong, W., Chen, Y., & Chen, L. (2025). SC-NBTI: A Smart Contract-Based Incentive Mechanism for Federated Knowledge Sharing. Sensors 2025, Vol. 25, 25(18). https://doi.org/10.3390/s25185802

Zheng, R., Sumper, A., Aragues-Penalba, M., & Galceran-Arellano, S. (2024). Advancing Power System Services With Privacy-Preserving Federated Learning Techniques: A Review. IEEE Access, 12, 76753–76780. https://doi.org/10.1109/ACCESS.2024.3407121