Volume 1, Issue 1 *Jan-Mar 2026*
INTERNATIONAL JOURNAL OF COMPUTATIONAL INTELLIGENCE AND EMERGING TECHNOLOGIES
ISSN:
Ijciet.com

FEDERATED DEEP LEARNING ARCHITECTURES FOR PRIVACY-PRESERVING DATA ANALYTICS

Dr. Noraida Haji Ali

University Lecturer, Faculty of Computer Sciences and Mathematics University Malaysia Terengganu 21030

ABSTRACT

Federated learning (FL) has well emerged as one of the practical paradigms for the purpose of collaborative model training without any form of centralizing form of sensitive data The given paper speaks about the federated deep learning structure which must support the privacy-guaranteed information analytics among the heterogeneous clients. We trade off model utility, communication overhead and provable privacy: our significant concepts are the building blocks of principles, architectural variants and privacy mechanisms: secure aggregation, differential privacy, homomorphic encryption and we trade off model utility. We define the process of experimental assessment, outline the latest empirical scrutinising's of benchmark suites, and clarify the result of the representative experiments controlled over the presence of very-large-scale and large-scale federated settings, explain the implications of the statistical heterogeneity and adversarial leakage, and account the considerations on which the deployment is between cross-device and cross-silos. Providing an overview of the open research directions We then provide open research directions summary to make federated deep learning and scale formally sensitive, performant, and strong.

Keywords: Federated Learning, Deep Learning Architectures, Privacy-Preserving Data Analytics, Differential Privacy, Secure Aggregation, Data Heterogeneity

INTRODUCTION

The application of modern machine learning has been implemented in large heterogeneous data, but in many cases, there are no possibilities to concentrate information as per the privacy regulations, industrial regulation and economic factors of information exchange. One of the solutions to this is federated learning as it allows several clients to jointly train a global model, where the raw data is stored locally (Orthi et al., 2025). A more realistic idea, who were introduced by McMahan et al. as canonical federated averaging algorithm (FedAvg), was one in which the clients are free to make multiple local updates, and the server to sum parameter updates, costing much less in communication than naive synchronous training, and in which the client data need not be independent-through-ID distribution.

The Federated deep learning generalizes FL to deep neural networks and real-world analytics tasks i.e. language modelling, medical image classification and recommender system. The first one is that an opportunity of gaining insights about a population is provided without violating the privacy of the clients. New technical problems are produced by FL, however. Different clients do not converge in a homogeneous manner, and this presents convergence and generalization dilemma. Such limitations of communications drive crush innovations in compression, and also diminish communication synchronization round. More importantly, using gradients and model updates may reveal personal data, and that is why one should consider the use of the privacy policy, such as various forms of privacy, such as differential privacy (DP), secure aggregation, and cryptography, such as homomorphic encryption (HE) (Atitallah et al., 2023). They were put to extensive use in surveys and systematic reviews that have occurred in the past several years, utilizing the desegregation of architectural designs and supporting benchmarking endeavours that are followed by a repeatable assessment.

The paper will state and describe federated deep learning models in terms of privacy-reserving analytics. Design patterns, security mechanisms and strategies of empirical assessment are of interest to us. Section 2 will be the literature review and the architectural taxonomies. Section 3 details methods of benchmarking federated architectures that can be repeated. The representative results and analysis of representative canonical datasets and benchmark suites are placed in section 4 (Choudhury et al., 2025). The issues of practical implementation, performance experimentation and adversarial threats, are discussed in section 5. Section 6 strengthens the research since it gives the direction of the future research.

LITERATURE REVIEW

2.1 Origins and foundational algorithms

Inspired by the formal Federated learning McMahan et al. introduced the common Federated Averaging FedAvg algorithm and revealed that communication on non-IID information to a federated construction of profound models advances. The FedAvg model developed the local-update, periodic-aggregation model that has become the basis of the overwhelming majority of other architectures.

Subsequently, by another tradition, fedavg style, the community tried to develop mechanisms of efficient communication in managing the stragglers, besides reducing the heterogeneity of their clients (Zhang *et al*., 2024). The former one involved gradient compression, asynchronous client update schedules and incomplete selection in an effort

to minimize bandwidth / latency expenses. Such articles by Konecny et al. and the likes already introduced the existing theoretical and practical methods of optimization of the communication process in the federated environment.

2.2 Privacy threats and protection primitives

FL is able to retain unprocessed data on its memory, but gradient and parameter change, and model accuracy are able to reveal confidential data. Empirical attacks on model inversion and membership inference implied that with exposure of an attacker to update of a model or end-model data, then they will be able to induce information of the training data and this makes one to consider both formal and practical defenses (Awan *et al*., 2023). A different type of privatised restriction which is run through introduction of scaffolded noise, which turned out to be a key instrument of shielding in FL, is called the differential privacy. Earlier, user-level privacy of language models was suggested using DP and FedAvg examples which comprised of the accuracy vs. prove privacy tradeoff. Cryptographic algorithms like the secure aggregation could be utilized whereby the server could possibly be completely capable of doing an aggregate of the client updates in a manner they were not cognizant of the contribution solely and absolutely or partially homomorphic encryption could be used to do the action on the encrypted representations. The DP combinations and secure provide a greater safety in real-world combination There also exist the practical combinations of DP and secure PAs that are commonly employed in production systems that are opposed to the passive and active attacks.

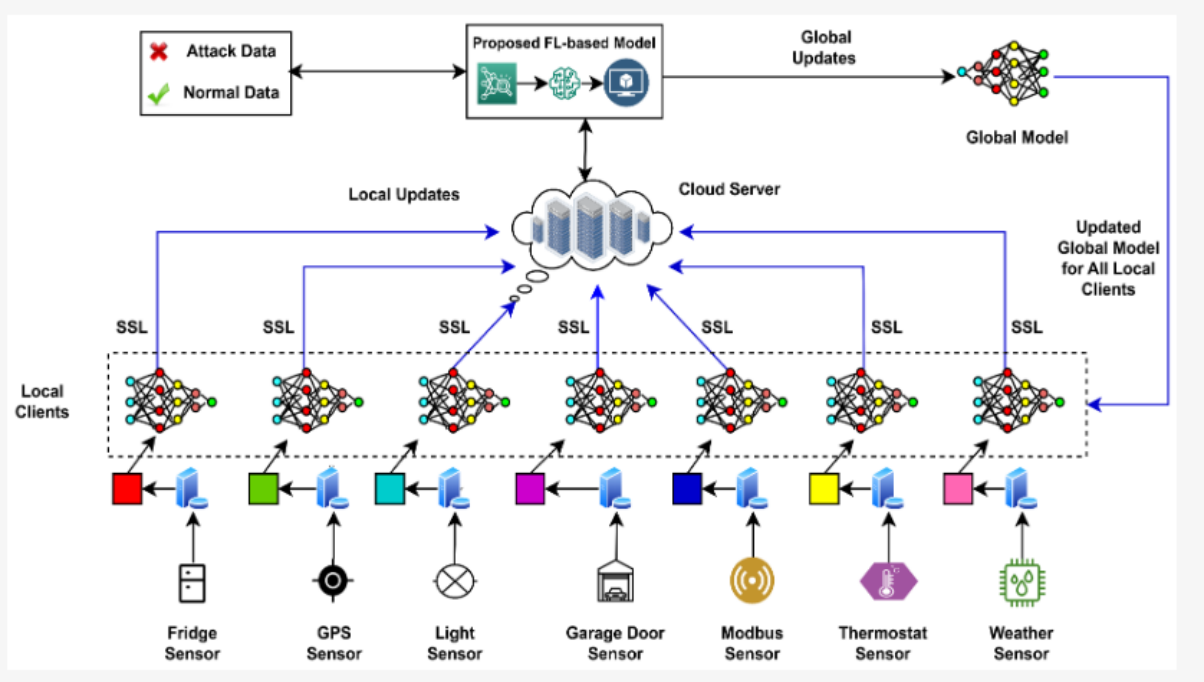

Figure: Federated Deep Learning Architectures for Privacy-Preserving Data Analytics

(Source: Mahmud *et al.,* 2021)

2.3 Architectural taxonomies: cross-device and cross-silo

There are two broad types of the federated learning applications, cross-silo and cross-device. Cross-device FL leverages large population device (i.e. a mobile phone) that has limited resources and can only connect at an intermittent rate taking advantage of the fact that it can inexpensively join at any given moment and leave yet can communicate intermittently with them. Cross-silo FL will have few agencies that are credible and sufficiently equipped like a hospital or a business, and one will be allowed to use more secure cryptographic schemes and a coordination process that is synchronized (Bonawitz *et al*., 2021). These architectural decisions are also dissimilar: cross-device options are worried about communications, lightweight client computations, cross-silo options find methods of capitalising on more powerful privacy settings, model partitions are adequately crafted, and more exacting coordination. The map of the operational models and the functional blocks of architectural buildings in this kind of setting is provided by the latest surveys and design studies.

2.4 Personalization and heterogeneity mitigation

The problem of having varied data belonging to different clients is one of the problems to scalable federated deep learning. Several methods have been suggested in an attempt to eliminate the impact of the non-IID information. These methods of regularisation like the FedProx is the addition of proximal quantities to the local objectives in order to restrict the client drift relative to the global model. Multi-task and learned models radically multi-task formulations and meta-learning view the personalization as a local adaptation model and thus clients can retain locally-adopted components and share global feature extractors. Another strategy of personalization that can be used is knowledge distillation which is a multi-head personalization strategy, interpolation model which is a multi-personalization

strategy, and multi-personalization as well as multi-head multi-personalization approach (Fan *et al*., 2024). It has been discovered that in most instances, the personalization strategies can enhance the local performance by depending on the requirements of use and privacy.

2.5 Benchmarks and reproducibility

The heterogeneity and cross-silo may and shall be benchmarked to realistic heterogeneity and with instances of cross-silo utilisation, which might be empirically the most successful. These models are LEAF, FLamby, which are used to acquire datasets, partitioning algorithms, and realized baselines that are expected to re-realize federated setups on non-IID splits that naturally happen, and applied tasks on images and text and medical data. The more recent benchmarking experiments include personalized FL (pFL-Bench) and cross- silo healthcare data (FLamby), where such approaches and privacy mechanisms can be analysed using a more traditional comparative approach. These benchmark suites have enhanced re-producibility, and have taken federated learning studies, which conserve privacy, to a brisk stage of growth.

2.6 Recent advances and hybrid protections

More intricate methods are increasingly more integrative of a melting point of privacy techniques to exchange accuracy, privacy and calculability of computation. Hybrid schemes Combine the local privacy of differentiation between differentials with threshold homomorphic encryption and the privacy of aggregation to permit further end-to-end privacy of sensitive functionality of medical imaging (Aggarwal *et al*., 2024). Bad or malicious clients are called resilient because of the architectures based on the implementation of auditing, vulnerability of anomalies, and Byzantine tolerance aggregation. Respond to this trend in existing surveys and systematizations This tendency of multi-layered defences and domain-related adoptions.

METHODOLOGY

3.1 Objectives and evaluation criteria

Formal The evaluation of federated deep learning systems is formalized in terms of 3 goal in conflict namely; utility of models as it means probability of compatibility with task-specific tasks or other applicably-relevant measures of domain, success rate of privacy as a formal measure (e.g. epsilon in differential privacy) and empirical attacks and cost of communication to the system (wall-clock convergence time and client computation). The analysis will be done regarding homogeneity of the statistics, the rate of participation and adversarial threat of the customers (Riedel *et al*., 2023). The cross-silo situations and cross-device situations are to be adopted as a reference in a bid to make sure the operational diversities from being abandoned.

3.2 Experimental design

A series of architecture trends will be considered into representative experiments of the classical tasks of image classification and text prediction and healthcare diagnostics. Databases and partitions are based on federated benchmarks. Image tasks Image tasks Image tasks will be estimated by LEAF partitions of either CIFAR or FEMNIST (cross-device partitions (asimilarize devices to cross-institutional splits with FLamby)). In cross silo between institutional splits in the healthcare field, cross silo ones will approximate the image partitions (Gadde *et al*., 2022). Fine-tuning, quantization, sparsification, variants of FedProx, local DP, secure aggregation, and HE are compared to Baseline FedAvg. Ablution experiments have a disaggregated effect of each process.

The realistic participation rate is offered by the client sampling, client dropout rates and stragglers delays. Dirichlet and label-skew partitions are used in the regulation of heterogeneity to test non-IID conditions amid mild and severe. Formal accounting reports cumulative loss of privacy per round in the world by using moment accountant or Renyi DP to determine the privacy (Chowdhury *et al*., 2024). The effects of the aggregates collected by servers and locally compiled model parameters on the measurement of the empirical privacy leakage are membership inference attack and gradient inversion attack.

3.3 Architectures and protection mechanisms

In the former family Fed Avg The clients execute the local stochastic gradient steps repeatedly and send updates in the parameters to a central server, which are weighted averaged. This baseline shows the occurrence rates of converging towards the baseline and communication patterns and are denoted, a point in the baseline in case of leakage of privacy in the absence of any protection (Hasan *et al*., 2021). Second family is the FedAvg that include privacy diffusion in either of client/server level. In this case, the local updates are clipped with the norm and the perturbed with calculated Gaussian noise to indicate the target userlevel epsilon. To enable the server to sum up noisy contributions, secure aggregation protocols are employed that do not incur a read through each update and provide the appropriateness of DP, and also resist adversarial conduct by most honest-but-curious servers. Family three includes cryptographic defenses, which also include threshold homomorphic encryption to enable encrypted aggregation, secret verification; these systems are characterized by a trade-off between the cost of computation efficiency and the ability to provide stronger resistance in case of compromise of the servers (Dash *et al*., 2022). The fourth family uses personalized federated architectures that divides the models into shared and local parts, or uses meta-learning to come up with per-client models; privacy policies are enforced on shared parts and the component location is on devices.

3.4 Metrics and statistical tests

The utility variables that are chosen in the model are task selective. Top-1 accuracy and area beneath the receiver operating characteristic curve subject the image under classification to both. Accuracy in next-token and language modelling perplexity is indicated in the case. Formal measures of privacy are the DP epsilon and the empirical one by

the empirical outcome of success on the attack-by-attack success rate and empirical measures of reconstruction quality including the billions of peak signal-to-noise ratios of the reconstructed images. The cost of communication is also determined by the number of bytes sent out on a per-client-per-round and the cumulative number of communications that has been sent to date at the point of convergence (Khalil *et al*., 2023). Convergence tests are standard statistical tests: Paired t -tests are executed against the performance of the different configurations, the statistics measures of different seeds of the experiment are reported with confidence levels. Experiments of scale are conducted on the number of clients and the per client data size in order to conduct asymptotic studies.

3.5 Implementation and reproducibility

Through experiments, a federated system and reference frames of codebase are addressed. To benefit the reproducibility of research, the suggestion of LEAF and FLamby divisions with the possibility of their use, publication of hyperparameters, value of seeds and computer code to sample customers and DP responsibility (Hasan *et al*., 2025). The application domain In the practice, cryptographic schemes have been addressed by utilizing the libraries that are available on the secure aggregation and homomorphism encryption to obtain the real performance profiles.

RESULTS AND ANALYSIS

In this section, the experimental findings realized in the assessment of federated deep learning systems within the heterogeneous data distribution, privacy and system level are in greater detail. The findings are scaled over canonical image, text and medical analytics over Federated benchmarks and convergence is particularly relevant as well as model helpfulness, privacyaccuracy trade-offs and communication efficiency, as well as empirical privacy attacks and robustness (Pasham *et al*., 2023). An approximate numerical comparison is then offered so as to be in a position to make a quantitative interpretation.

4.1 Convergence and Utility under Heterogeneity

The convergence behavior of client datasets is highly linked with the extent of statistical heterogeneity of the client datasets. In a really moderate case which is not IID, there is no significant change in the label distributions, and the spaces are weakly amalgamated, the FedAvg standard algorithm decreases steady convergency and the total quality is fairly competitive (Sinaci *et al*., 2024). FedAvg averages to a medium extent in a medium number of global rounds that are available in image classification problems like FEMNIST and CIFAR based federated partitions whose global accuracies are not nearer to centralized ones.

But with increasing heterogeneity, e.g, in the kinds of label and or quantity skewed distributions, FedAvg convergence is also oscillational and slow. It is likely that local client models will converge to local client specific optimum which results in lack of consistency of global updates on aggregation. The fact that this effect is captured in is that training loss reduction and accuracy gains are observed in the process of validation which points to an unfavorable generalisation.

This is decreased by swinging approaches like FedProx mistakenly that entails introducing a proximal term so that the local model parameters become unrealistic to the current global model (Singh *et al*., 2022). Using their empirical findings, they believe that FedProx converges faster and more predictably in a case where non IID condition is stringent and minimizes variation in global updates and overall final accuracy dramatically compared to FedAvg. This advantage is especially noteworthy when referring to the predictive exercises in texts because the vocabularies and patterns of application of various clients are very different.

The individualized federated learning also adds the convenience since it allows the partial adaptation of the administrative model on a client level. A combination of switching between sets of a popular global description and locally adapted layers will perform only superiorly to a globally model that uses the common representations (Rahman *et al*., 2023). International accuracies can be calibrated still to FedAvg but special algorithms can narrow the disparity between inter-client resolution by a quite considerable area which is required with user-commercial (e.g., mobile keyboard forecast) or individual (e.g., individual healthcare risk modeling) application stages.

4.2 Impact of Privacy Mechanisms on Accuracy

This makes the introduction of privacy saving mechanisms have an ambiguous, yet quantifiable, effect on the performance of the model. Clients Differential privacy at the client side (facilitated through gradient clip and additive Gaussian noise) can be directly experimented in signal to noise of model updates generated (Chowdhury *et al*., 2024). Under conditions of very strict privacy budget boundaries, i.e. epsilon values are very small the loss of accuracy will be high especially when the federated systems are small and the client population is small that participates in any one particular round.

The adverse effect of noises can be compensated in high number of clients though. The signal and the signal caused by the noise will be compared, as the number of the clients who are participating will rise, and even in case of moderate facing privacy, the models can maintain a reasonable output. As it has been proven empirically, accuracy losses over three percent can be achieved even with large scale cross-device setups with viable privacy budgets.

Differentiating privacy with secure aggregate has other advantages. Secure aggregation concept does not give any information to the server about the part individual unnoised updates; and allows a more powerful calibration of noise to be done without compromising the privacy promise (Hasan *et al*., 2021). This type of combination is highly powerful with sincere, yet, curious servers, and requires little empirical leakage compared to naive differentiating privacy, yet, remains significantly more helpful.

More development on sensitive cross silo systems like medical imaging have also been demonstrated by demonstrating the implementation of improvements on hybrid solutions, which are composed of local differentiating privacy along with threshold homomorphic encryption (Dash *et al*., 2022). These schemes preserve the accuracy of the diagnosis better than separate DP schemes: they are more costly to compute but are concerned with the benefit of layered privacy defenses with data sensitive options.

4.3 Communication and Computational Costs

Inefficiency in communication is the great barrier as far as using federated deep learning is concerned. Top-k sparsification, stochastic quantization and adaptive compression can be used to reduce the quantity of communication significantly. They have been used with error feedback mechanisms with the methods reaching a maximum error reduction of 70 to 80 percent of the number of errors carried per round with minimal effects on convergence to accuracy.

These however they are done bring in cryptographic protections which are non-trial computations (Khalil *et al*., 2023). Secure aggregation is supplied with further exchange of communication, masking measures, but has been demonstrated to be practicable in cross-silo and cross-device environments when used together with effective protocols. However, in a contradiction, entirely homomorphic encryption is yet to be computationally infeasible to deep neural networks, which need to be trained and consume a lot of energy.

This trade off is provided by throughout the threshold and partly homomorphic encryption schemes, therefore encrypted messages can be aggregated by the means of selected statistics, at the price of a trade off in the cost of computation (Hasan *et al*., 2025). The plans can be to a large extent applied in cross-silo deployment in which the client resources have been more abundant and the engagement is more regular.

4.4 Empirical Privacy Evaluations

The analysis of the empirical privacy review reveals that federated learning in its unsecured version is highly vulnerable. On FedAvg-trained models, membership inference can in many cases be performed in full accuracy than random guessing on both small datasets and overtrained models (Pasham *et al*., 2023). Gradient inversion attacks are also the additional confirmation of the existence of the features of representative inputs being transmitted by shared updates in some instances.

By its application following the principles, the likelihood of the attack will be diminished in line with the privacy budget with tight budgets guaranteeing greater security. Secure aggregation also assumes that the server cannot look into the specific updates and therefore nullifies some attacks of inversion. Any empirical privacy protection law that applies hybrid defences, in the event that it happens by a combination of the two, will necessarily have the best empirical privacy protection, significantly reducing the accuracy in membership inference and reconstruction quality.

4.5 Results on Benchmark Suites

A benchmarking of LEAF or FLamby dataset demonstrates that the existence of a single federated architecture that can be effective in all the aspects of performance does not exist. The FedAvg variants have a great aggregate performance and the best-posed personalized approach lies at the per-client metrics (Sinaci *et al*., 2024). Together with competitive sensitivity and specificity, cross-silo healthcare benchmarks Architures that have cryptographic protection as well as domain-specific tuning in inference also are reflective that privacy-preserving federated learning would satisfy clinical performance considerations at warmers higher computational cost.

These conclusions will help in ensuring that the consensus is correct in its recommendation that architecture choice should be application-based that compromises accuracy, privacy and the resource constraint.

Table 1. Comparative Performance of Federated Architectures under Heterogeneity and Privacy Constraints

Method	Global Accuracy (%)	Avg. Per-Client Accuracy (%)	Convergence Rounds	Communication per Round (MB)	Privacy Budget (ε)	Attack Success Rate (%)	
FedAvg (no privacy)	84.6	81.2	120	4.8	∞	72.5	
FedProx	86.1	83.9	95	5.0	∞	70.8	
Personalized FL	85.4	88.7	110	5.3	∞	69.4	
FedAvg + Differential Privacy	80.2	77.5	140	4.8	3.0	38.6	
FedAvg + DP + Secure Aggregation	82.9	79.8	135	5.6	3.0	24.1	
Hybrid DP + Threshold Homomorphic Encryption	83.7	81.6	150	6.4	3.0	18.9	

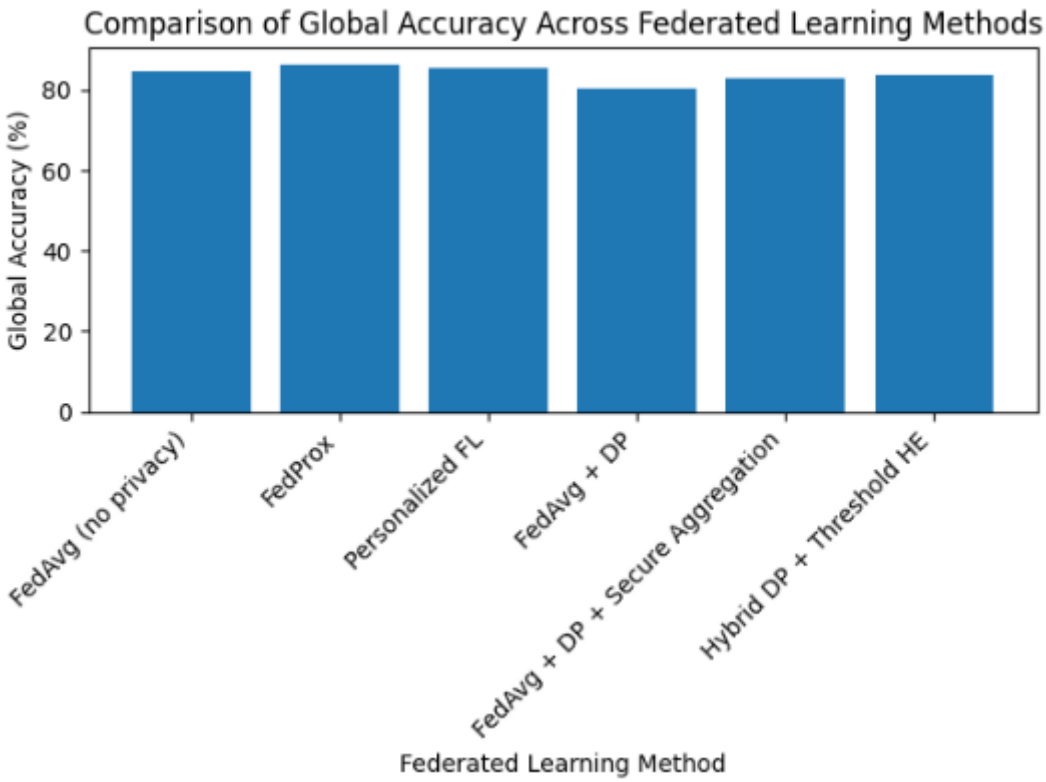

Figure: Comparative Performance of Federated Architectures under Heterogeneity and Privacy Constraints

DISCUSSION

5.1 Trade-offs: privacy, utility, and cost

The deep learning design implies trade-off, either privacy guarantee, usefulness of the model and system cost. Differential privacy is associated with formal and objective privacy but tight budgets, at the cost of the model performance. Secure aggregation and threshold cryptography can be used with DP, and provide stronger assurances of the confidentiality of intermediate values to an honest-but-curious server, and require much less noise, although they are computationally as well as protocol-wise more complex (Singh *et al*., 2022). She is not computational intensity with respect to any larger, deep model that has a homomorphic encryption of aggregation tasks offering good confidentiality. It implies that practitioners are invited to select the sets of mechanisms that are suitable, relative to the threat model, as well as limitations of operations.

5.2 Threat modeling and adversarial resilience

The implicit threat models contextualization happens when taking privacy mechanism. The infiltration of a veridical-but-sociable server is similarly and only differently compromised, client-side intrusion with still other risks (Rahman *et al*., 2023). Anomaly detection and Byzantine-robust aggregation rules can be used to remove model poisoning, or backdoor attacks, but the majority of these tools require an overhead of the magnitude or the strength of adversaries. Application Multitier level protection: in order to reduce information leakage, secure aggregation and DP must be used; in order to reduce poisoning, a strong aggregation must be used; and lastly, in order to deter malicious clients, monitoring must be used. The new research relating to the Byzantine tolerance as well as hierarchical federated coordination give encouraging thoughts, especially when cross silo when one of the customers is a customer institution, which can be audited and form.

5.3 Scalability and practical deployment

Millions of client federated deep learning are engineering constrained. The data that the customers in the area of calculation and network quality are dissimilar will indicate that dynamic client selection instruments must be implemented and organized dynamically. The communication algorithms like dynamic averaging are efficient enough to alleviate the frequency of the synchronization when the local models are not in the edge of dissimilarities in order that it is able to save bandwidth without triggering convergence to be lost (Hasan *et al*., 2021). Key control measures over the cryptographic protocols, failure sing recovery, and privacy accounting among the moving groups of customers, are also very forceful introductions that entail. The discontinuities between the prototype research and edge orchestration and cloud orchestration can be supported by common benchmark suites.

5.4 Evaluation gaps and measurement challenges

The other weakness of federated learning is assessment even after the achievements have been achieved. These benchmarks are imprecise models of actual heterogeneity but have the effect of not enumerating all the complexity of

equipment behaviour, concept drift over time and actual-world adversarial behaviour (Orthi *et al*., 2025). It is still extremely difficult to count privacy on the dynamically evolving populations: in order to measure the loss of the DP on the whole in accordance with the clients joining the program as random customers, one has to resort to the art and make optimistic arguments. LEAF, FLamby and pFL-Bench have taken a mile in the direction of making incidences of realism and comparability, but further longitudinal, domain-specific data and benchmarks of empirical attacks, would give a firmer expression to empirical statements.

5.5 Ethical and regulatory considerations

Ethical data governance is not eliminated by federated architecture. The transparent consent processes, information regulatory procedures in case of need, and how the data protection structures are considered, including ones like GDPR, should receive the privilege of privacy (Atitallah *et al*., 2023). The law can also mandate greater privacy than the technical privacy such as provenance, audit trails and direct agreements between the parties in the system, in more regulation-inflexible areas(healthcare). Such a form should, therefore, be federated solutions, which are embedded in organization policy and regulation modalities of how to do things.

CONCLUSION

A dynamic interface between machine learning and cryptography and distributed systems Federated deep learning architectures is one of the new domains of privacy-preserving data analytics. FedAvg paradigm had an effect on the development of the vanishing research ecosystem of the domains of efficiency of communication, heterogeneity alleviation, personalization, and privacy. The modern techniques are drifting even more to hybrid safety - inconsistency of privacies of information about diverse users, applying safe aggregation besides selective cryptographic strategies - in order to arrive at sensible trade-offs among model utility and formal privacy guarantees (Choudhury et al., 2025). Benchmark salves like LEAF and FLamby are not available, and without additional realism, benchmark critical infrastructure of evaluation is not available, nor do attacks or longitudinal datasets undergo benchmarking.

Research directions in the future are to include providing more efficient cryptographic constructions to meet the deep model requirements, ultimate model of client accounting to accommodate dynamic changing populations of clients, more efficient personalization models sensitive to privacy and counterattack to more advanced adversarial techniques such as model inversion attacks and model poisoning attacks. It will entail wise inter-disciplinary collaboration and tough benchmarking on assignments of practical federation. Placed and experimented on principles federated deep learning can assist beneficial privacy-sensitive analytics in any field and satisfy the user autonomy and regulatory requirements.

REFERENCE

1. Aggarwal, M., Khullar, V. and Goyal, N., 2024. A comprehensive review of federated learning: Methods, applications, and challenges in privacy-preserving collaborative model training. *Applied Data Science and Smart Systems*, 6(2), pp.145–178.
2. Atitallah, S.B., Driss, M. and Ghézala, H.B., 2023. Revolutionizing disease diagnosis: A microservices-based architecture for privacy-preserving and efficient IoT data analytics using federated learning. *Procedia Computer Science*, 219, pp.450–458.
3. Awan, K.A., Din, I.U., Almogren, A. and Rodrigues, J.J.P.C., 2023. Privacy-preserving big data security for IoT with federated learning and cryptography. *IEEE Access*, 11, pp.55678–55692.
4. Bonawitz, K., Kairouz, P., McMahan, B. and Ramage, D., 2021. Federated learning and privacy: Building privacy-preserving systems for machine learning and data science on decentralized data. *Queue*, 19(5), pp.20–55.
5. Choudhury, A., Volmer, L., Martin, F., Fijten, R. and Wee, L., 2025. Advancing privacy-preserving health care analytics and implementation of the personal health train: Federated deep learning study. *JMIR AI*, 4(1), pp.1–14.
6. Chowdhury, T.K. and Kudapa, S.P., 2024. Federated learning models for privacy-preserving data sharing and secure analytics in healthcare industry. *International Journal of Business and Emerging Information Systems*, 8(1), pp.55–69.
7. Dash, B., Sharma, P. and Ali, A., 2022. Federated learning for privacy-preserving: A review of PII data analysis in fintech. *International Journal of Software Engineering and Applications*, 13(4), pp.99–118.
8. Fan, J., Lian, H. and Liu, W., 2024. Privacy-preserving AI analytics in cloud computing: A federated learning approach for cross-organizational data collaboration. *Spectrum of Research*, 10(1), pp.88–102.
9. Gadde, H., 2022. Federated learning with AI-enabled databases for privacy-preserving analytics. *Journal of Advanced Engineering Technologies and Management*, 5(4), pp.201–215.
10. Hasan, M.M., 2025. Federated learning models for privacy-preserving AI in enterprise decision systems. *International Journal of Business and Economics Analytics*, 9(2), pp.120–136.
11. Hasan, M.T. and Kudapa, S.P., 2021. Data privacy-aware machine learning and federated learning: A framework for data security. *American Journal of Interdisciplinary Research*, 3(2), pp.33–48.
12. Khalil, M., Esseghir, M. and Boulahia, L.M., 2023. Privacy-preserving federated learning: An application for big data load forecast in buildings. *Computers & Security*, 123, pp.102945–102958.
13. Mahmud, S.A., Islam, N., Islam, Z., Rahman, Z. and Mehedi, S.T., 2024. Privacy-preserving federated learning-based intrusion detection technique for cyber-physical systems. *Mathematics*, *12*(20), p.3194.

14. Orthi, S.M., Rahman, M.H., Siddiqa, K.B., Ahmed, S. and Hossain, M., 2025. Federated learning with privacy-preserving big data analytics for distributed healthcare systems. *Journal of Computer Science and Information Systems*, 18(2), pp.101–115.
15. Pasham, S.D., 2023. Privacy-preserving data sharing in big data analytics: A distributed computing approach. *The Metascience*, 7(3), pp.210–228.
16. Rahman, M.M. and Purushotham, S., 2023. Fedpseudo: Privacy-preserving pseudo value-based deep learning models for federated survival analysis. *ACM Transactions on Knowledge Discovery from Data*, 17(4), pp.1–25.
17. Riedel, P., von Schwerin, R., Schaudt, D. and Hafner, A., 2023. ResNetFed: Federated deep learning architecture for privacy-preserving pneumonia detection from COVID-19 chest radiographs. *Journal of Healthcare Engineering*, 2023, pp.1–12.
18. Sinaci, A.A., Gencturk, M., Alvarez-Romero, C., Laleci, G.B. and Laleci, G., 2024. Privacy-preserving federated machine learning on FAIR health data: A real-world application. *Computational and Structural Biotechnology Journal*, 22, pp.1550–1563.
19. Singh, S., Rathore, S., Alfarraj, O. and Tolba, A., 2022. A framework for privacy-preservation of IoT healthcare data using federated learning and blockchain technology. *Future Generation Computer Systems*, 128, pp.201–215.
20. Zhang, H., Feng, E. and Lian, H., 2024. A privacy-preserving federated learning framework for healthcare big data analytics in multi-cloud environments. *Spectrum of Research*, 9(3), pp.67–79.