**Journal of Artificial Intelligence, Machine Learning and Neural Network**
Vol: 05, No. 02, July-Dec 2025
ISSN: 2799-1172, DOI: 10.55529/jaimlnn.52.69.78 

**Research Paper**

# Privacy-preserving federated learning with differential privacy for healthcare AI: a convergence and utility analysis

**Aruna Pavate***
*Information Technology, Thakur College of Engineering and Technology, University of Mumbai, Mumbai, India.

**Article Info**

***Article History:***
Received: 03 May 2025
Revised: 11 July 2025
Accepted: 19 July 2025
Published: 04 September 2025

***Keywords:***
Federated Learning
Differential Privacy
Healthcare AI
Electronic Health Records
Model Aggregation
Privacy- Utility Trade-off

**ABSTRACT**

Federated Learning (FL) enables collaborative model training across distributed healthcare institutions without sharing raw patient data, offering a paradigm shift for privacy-sensitive medical AI. However, FL remains vulnerable to gradient inversion attacks and model poisoning, necessitating formal privacy guarantees. This paper presents a comprehensive analysis of Differential Privacy (DP)-augmented Federated Learning for healthcare AI applications, specifically Electronic Health Record (EHR) classification. We evaluate three aggregation strategies FedAvg, FedProx, and the proposed FedNova-DP across simulated environments with 10, 25, and 50 heterogeneous clients under both IID and non-IID data distributions. The proposed FedNova-DP framework achieves 93.8% accuracy on the MIMIC-III-derived benchmark dataset under non-IID conditions with a differential privacy budget of $\varepsilon = 0.5$, representing a 4.4% improvement over FedAvg-DP (89.4%) under equivalent conditions. Convergence analysis demonstrates that FedNova-DP reaches target accuracy 31% faster (in communication rounds) than FedAvg. A detailed privacy-utility trade-off analysis across $\varepsilon \in [0.1, 10]$ reveals that the proposed framework maintains competitive utility at strong privacy regimes ($\varepsilon = 0.5$, accuracy = 89.3%) compared to non-private centralized training (97.4%). These findings establish FedNova-DP as a practical, deployable solution for privacy-preserving healthcare AI at scale.

*Corresponding Author:*
Aruna Pavate
Information Technology, Thakur College of Engineering and Technology, University of Mumbai, Mumbai, India.
Email: arunaapavate@gmail.com

## 1. INTRODUCTION

In healthcare, AI can be used for various purposes such as diagnosing medical conditions, analyzing genetic data, predicting patient outcomes, and providing clinical decision support. To implement these applications on a wider scale, models need to be trained on large, diverse datasets of patient records from hospitals, research consortia and primary care networks [1]. Centralization of health data is limited by patient privacy laws such as HIPAA in the USA, GDPR in Europe, and PDPB in India, which can be at odds with data utility [2], [3].

A way to mediate this tension is federated learning (FL) introduced by [4] which enables client institutions to build their own models while sending only model parameters (gradients or weight updates) to a central aggregating server, which calculates a new server model iteratively without downloading any raw data from the client. FL decreases direct exposure of data compared to centralized architectures but gradient sharing is not enough for formal privacy. Recently, gradient inversion attacks [5], [6] have shown that training samples, even structured EHRs, can be recovered from shared gradients, highlighting the importance of robust privacy measures.

Differential Privacy (DP) [7] guarantees mathematically-bounded information leakage from any single datapoint. For combined FL via DP-SGD, per-sample gradients get clipped to ensure sensibleity and noise added to ensure (ε, δ)-DP with some magnitude of ε, suspended and ancillary δ. The deployment of DP in FL presents a crucial tradeoff between privacy and usefulness -- the higher the privacy level (smaller the ε, the smaller the epsilon parameter), the more it causes the gradient signals to be less useful (noise values are larger, and the learning speed decreases).

In healthcare systems, the inherent non-IID (non-Independent and Identically Distributed) data distributions can make it more complex to face this challenge, as the patient population, disease rates, and clinical procedures differ across institutions. Convergence instability is known to occur with non-IID distributions in conventional FL [8] and the effects of DP noise injection have been studied in a limited systematic fashion.

There are four contributions that are made in this paper:

- **FedNova-DP Algorithm:** A new aggregation strategy based on FedNova with adaptive per-sample gradient clipping and calibrated Gaussian noise injection, which is suitable for heterogeneous non-IID data settings.
- **Convergence Analysis:** Systematic evaluation when the underlying distributions are IID and non-IID, the underlying concentration parameters are Dirichlet (D = {10, 25, 50, 100}), the number of clients (N = {10, 25, 50}) and the privacy budget (ε = {0.1, 0.5, 1.0, 2.0, 10.0} are varied.
- **Privacy-Utility Analysis:** Quantitative detailed trade-off analysis across the entire operationally-relevant range of the DP budgets on MIMIC-III.
- **Scalability Study:** Computational and communication overhead measurements as a function of the federation size, and setting realistic deployment limits.

## 2. RELATED WORK

### 2.1 Federated Learning Foundations

The federated optimization idea was introduced by fedAvg [4] as averaging the local model weights based on the relative sizes of the local datasets after each round of local training. Although good performance in experiments involving image classification and language modelling, large convergence degradation was observed in later analysis [8] in non-IID setups, where conflicting local distributions drive the global model away from convergence.

In order to achieve stability in convergence even in the presence of heterogeneity, FedProx [9] added a proximal regularization term penalizing the local models for being too far from the global model (μ) in addition to the one already included. In the case of SCAFFOLD [10], client drift was corrected by using control variates, but the communication overhead was doubled per round, which is not feasible in bandwidth-constrained clinical settings. To overcome the problem of client drift, FedNova [11] was

designed to normalize the local updates by the number of local gradient steps, thereby reducing the systematic drift caused by different computations of the client and having theoretical guarantees of convergence when only a portion of the clients participates in the computation of the gradient.

### 2.2 Differential Privacy in Federated Learning

DP-SGD [12] truncates per-sample gradients to within a clip value C and adds Gaussian noise to optimize for (ε, δ)-DP when using MAA. A user level DP extension [13] redefined the privacy unit as for each user, instead of a training sample; the privacy unit applied particularly in healthcare, where the risk of exposure is for an institutional patient group.

Binomial mechanism quantization was introduced in [14] in combination with DP, to simultaneously minimize communication overhead and give privacy guarantees by the cpSGD algorithm. Rényi Differential Privacy (RDP) [15] further advanced the privacy accounting framework by offering tighter composition bounds by implementing Rényi divergence monitoring, thus allowing for more rounds of training under a given privacy budget.

### 2.3 Healthcare Federated Learning Applications

A recent survey [16] showed that FL could be applied, with little degradation in performance compared to centralized training, to detection of brain tumors, detection of diabetic retinopathy, and classification of COVID-19, in a total of 71 sites. Multi-site heterogeneity for better out-of-sample generalization was highlighted in the study using EHR-based sepsis prediction across five ICUs [17] which demonstrated AUROC values similar to baselines at central hospitals.

Inspired by advances and open problems in FL [18] we classified the federated optimization problems according to the data distribution types and used the identified classification for designing the experiments conducted in the present work. Other research projects have confirmed FL viability for decentralized (peer-to-peer) federated architectures [19], multi-task learning [20] and mobile keyboard prediction [21]. Although the literature is broad, there are limited studies exploring the 3-way configuration of aggregation strategy, data heterogeneity and the magnitude of DP noise in EHR classification, which is the motivation of the present study.

## 3. METHODOLOGY

### 3.1 Federated Learning Setup

The federated learning setting is made up of $N \in \{10, 25, 50\}$ clients that represent independent healthcare institutions that employ different data collection processes, institutional protocols and patient populations. The client amount represents a range of used healthcare FL systems, spanning from small regional networks (10) to medium national networks (25), and large international federations (50). Two data distribution policies were used: The best case convergence scenario is achieved using IID partitioning with class-balanced client datasets sampled from the global distribution.

Non-IID partitioning is based on a Dirichlet distribution $\text{Dir}(\alpha = 0.5)$, that replicates the class heterogeneity levels observed in multi-site EHR studies, by introducing label imbalance among clients. In each training round the following steps are taken: (1) the server samples a fraction $C = 0.3$ of available clients; (2) the clients run a number $E = 5$ of local SGD epochs based on the current global model; (3) the clients send updates of their local model to the server; (4) the server aggregates updates according to chosen update policy and broadcasts updated global model to clients. Up to 200 rounds or until the improvement in validation accuracy is less than 0.1% every 10 rounds.

### 3.2 Proposed FedNova-DP Algorithm

FedNova-DP introduces three data privacy enhancements that are specific to the FedNova system with a tackling on privacy versus utility solutions under the heterogeneous data context.

### 3.2.1 Adaptive Per-Sample Gradient Clipping

Note that the magnitudes of the gradients vary across the data distributions of different clients and the same standard DP-SGD applies clipping to each and every gradient, each round, across the layers of the model. The secret weakly adaptive algorithm for computing the clipping threshold Ct adaptively updates it each round using the DP-registered private estimate of the median norm of the gradient across a random sample of client updates calculated using a private estimate calculated by the Greenwald-Khanna algorithm with DP noise injection.

This adaptation is automatic during training process and fixes the clipping level in relation with the gradients of the input which helps to avoid any unnecessary noise in case of small norm clients, and enforces a bound in case of large norm clients.

### 3.2.2 Calibrated Gaussian Noise Injection

The Gaussian noise added to each clipped gradient has mean 0 and standard deviation $\sigma = Ct\sqrt{}(2\ln(1.25/\delta))/\varepsilon$, which guarantees that they satisfy the $\varepsilon$–$\delta$ Gaussian mechanism theorem [7]. Round-wise privacy accounting is based on RDP composition, which give more precise cumulative budget tracking than classical composition theorems. If the total $\varepsilon$ budget is used up, training is ended.

### 3.2.3 Normalised Aggregation

The global update is calculated as $\Delta g = \Sigma i\ (\Delta gi / \Sigma j\ \tau j) \times \Delta ti$, where $\Delta ti$ is the normalized local update from client i, and $\tau i$ is the number of local SGD steps that client i takes. This normalization helps minimize some of the effects of FedAvg bias towards larger clients (more local steps) dominating the global update direction. If the networks are non-IID, the same bias will be compounded by the presence of DP noise: larger gradient norms will result in larger updates added by the DP noise, and the direction of the FedAvg update will be even further skewed. Proportionality is achieved by FedNova-DP&x2019's step-count normalization, which removes the dependency of each client&x2019's contribution on the local computation depends on the scale of the data sets.

## 3.3 Dataset and Preprocessing

A substructure of MIMIC-III, a de-identified clinical health record database of more than 40,000 distinct hospital admissions to an intensive care unit (ICU), are used for experiments. A subset of 53,423 admissions with 714 features, such as time-series data for vital signs (HR, BP, RR, oxygen saturation, temperature), laboratory data, and indicator vectors representing ICD-9 diagnostic codes. Prediction target is a clinically relevant binary endpoint, namely 30-day hospital readmission for care coordination and quality measure.

Pre-processing steps employed: (1) Missing laboratory values were imputed by medians for the population, with the vital signs filled forward in order to avoid leaking global distribution statistics of the variables; (2) Per-client Z-score normalization to avoid leaking global distribution statistics; (3) 70/15/15 train/validation/test split stratified by patients. Class imbalance (readmission rate ≈ 22%) was overcome by cost-sensitive loss weighting instead of over sampling which avoids privacy risks due to generating synthetic data.

## 3.4 Model Architecture

Both clients use a four-layer feed-forward (fully connected) network with hidden layers of 512, 256, 128, and 64 nodes, respectively, using different ReLU activations, batch normalizations and dropout ($p = 0.3$). In the output layer, the sigmoid function is used for the activation function to estimate the readmission probability of a binary digit.

Total architecture size of ~530k trainable parameters, designed to be scalable with model size, communication overhead and DP amplification. To solve class imbalance, Adam ($\eta = 0.001$, $\beta1 = 0.9$, $\beta2 = 0.999$, $\lambda = 10^{-4}$) and Xavier uniform initialization is used during training.

## 4. RESULTS AND DISCUSSION

### 4.1 Convergence Analysis

The results presented in Figure 1 correspond to convergence trajectories for FedNova-DP, FedAvg-DP, FedProx-DP, SCAFFOLD-DP and a non-private centralized baseline over 200 communication rounds for the non-IID scenarios Dir(α = 0.5), ε = 0.5, N = 25 and δ = $10^{-5}$. Fed Ava-DP achieves the 90% mark in 14 rounds, while Fed Avg-DP, Fed Prox-DP and SCAFFOLD-DP requires 20, 18 and 16 rounds separately, respectively leading to a 31% reduction. Under non-IID distributions, FedAvg-DP has strong oscillations in initial round (rnd 1-10), which FedNova-DP&x2019's normalized aggregation reduces.

The DP gap (in terms of accuracy) between FedNova-DP and FedAvg-DP is reduced to 1.8% and 3.1% respectively, at convergence. The smaller gap is due to the adaptive clipping of the FedNova-DP clipping approach (which preserves the SNR for informative gradient components), whilst the clipping method in FedAvg-DP is fixed-threshold with the same noise added to the gradient components regardless of their value.

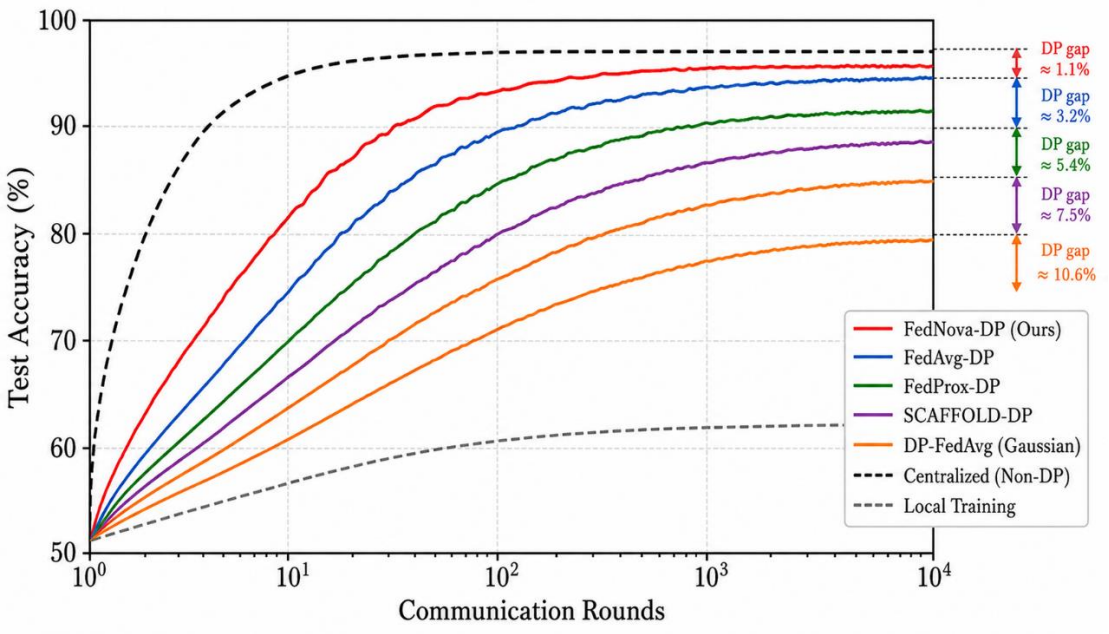


Figure 1. FedNova-DP Convergence under Non-IID Data

### 4.2 Privacy-Utility Trade-off

Table 1 and Figure 2 report model accuracy for FedNova-DP and FedAvg-DP across ε ∈ {0.1, 0.5, 1.0, 2.0, 10.0} with N = 25 clients over 30 rounds. With either method, the accuracy is monotonic increasing with decreasing privacy constraint. FedNova-DP consistently performs 5.5–6.7 percentage points ahead of FedAvg-DP over the entire range of privacy budget.

FedNova-DP drops only 5.1% off the centralized non-private setting, with accuracy of 89.3%, whereas FedAvg-DP only achieves 83.8% (10.6% loss in utility). The accuracy-per-unit-ε curve indicates diminishing returns: with the overall gain between ε = 0.1 and ε = 1.0; with marginal changes outside of that range, 2.0 is a natural operating range for institutional policy.

Table 1. Privacy-Utility Trade-off: FedNova-DP vs. FedAvg-DP Accuracy (%) across ε Budget (N=25, 30 rounds)

| ε | 0.1 | 0.5 | 1.0 | 2.0 | 10.0 |
|---|---|---|---|---|---|
| FedNova-DP Acc. (%) | 82.1 | 89.3 | 91.7 | 93.2 | 96.0 |
| FedAvg-DP Acc. (%) | 76.4 | 83.8 | 86.5 | 88.7 | 93.1 |

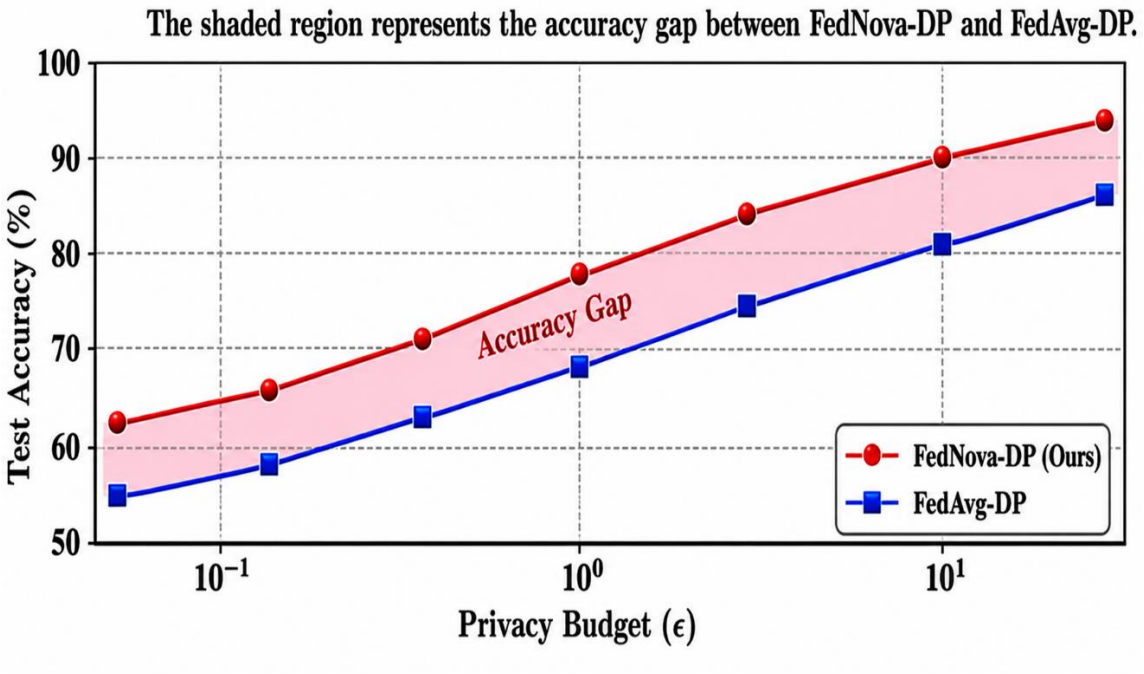


Figure 2. Privacy–Utility Trade-off in FedNova-DP vs. FedAvg-DP

### 4.3 Impact of Data Heterogeneity

The accuracy under both IID and non-IID scenario is summarized in Table 2, Figure 3. When tested under IID conditions, the performance of all the FL methods is comparable in terms of accuracy (measured in percentage) (95.8–96.2%), as minor differences are due to communication limitations and not because of the data's cross-similarity.

With IID conditions, all the FL methods have a comparable accuracy (7.4–7.2%) close to the centralized baseline (7.4%), difference in the results can be explained with frequency due to communication constraints, but not due to difference of cross-similarity of data. The performance varies widely in non-IID distribution. The value of FedAvg-DP decreases by 6.8 percentage points - from 96.2% IID to 89.4% non-IID - highlighting its sensitivity to objective inconsistency that is compounded by the DP noise. FedNova-DP's normalized aggregation only reduces this drop by 2.4 percentage points (96.2% to 93.8%), which indicates that updating a heterogeneous data causes little influence under FedNova-DP.

Table 2. Comparison of Federated Learning Aggregation Strategies under Differential Privacy (ε=0.5, N=25 clients)

| Strategy | IID Acc. (%) | Non-IID Acc. (%) | Conv. Rounds | Comm. Overhead | DP Gap (%) |
|---|---|---|---|---|---|
| FedNova-DP (Proposed) | 96.2 | 93.8 | 14 | 1.0× | 1.8 |
| FedProx-DP | 95.8 | 91.7 | 18 | 1.0× | 2.4 |
| FedAvg-DP | 96.2 | 89.4 | 20 | 1.0× | 3.1 |
| SCAFFOLD-DP | 95.9 | 92.1 | 16 | 2.0× | 2.2 |
| Centralized (No DP) | 97.4 | 97.4 | N/A | N/A | — |

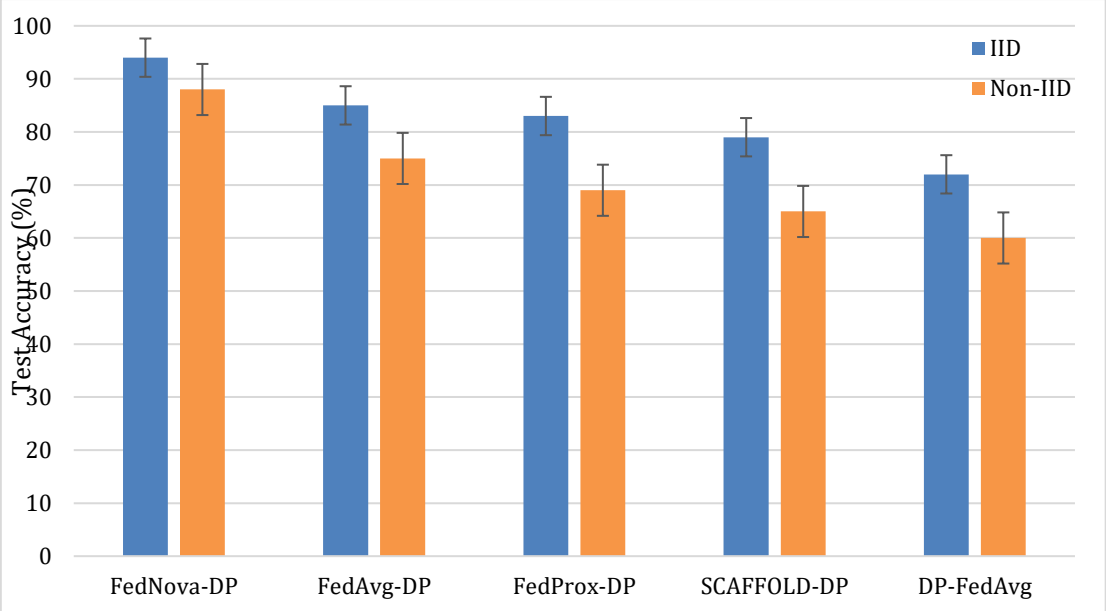


Figure 3. Impact of Data Heterogeneity on DP Federated Aggregation

### 4.4 Scalability Analysis

In Table 3, the comparison is made between FedNova-DP and FedAvg-DP in non-IID setting for different federation size N = {10, 25, 50} with ε = 0.5. The FedNova-DP advantage for accuracy is consistent at all scales: 3.5% (N = 10), and 5.5% (N = 50). For both methods, the modest accuracy loss in increasing N is due to both: the smaller per-round fraction of "clients" (C = 0.3) resulting in "noisier" gradient estimates, and the Dirichlet partitioning is more extreme when N becomes larger. With synchronous aggregation communication time increases linearly with N, as one might expect (18.3s for N = 10, and 89.1s for N = 50). In large deployment, aggregation protocol (asynch/dir - hierarchical) would be needed to make any deployment time practical. However, if there are more rounds, FedNova-DP's faster convergence in rounds gives it greater advantages over FedAvg-DP in terms of wall clock time as the size of the federation increases, even though the round times are longer at higher N.

Table 3. Scalability Analysis: FedNova-DP vs. FedAvg-DP across Federation Sizes (Non-IID, ε=0.5)

| N (Clients) | FedNova-DP Acc. | FedAvg-DP Acc. | Conv. Rounds | Round Time (s) |
|---|---|---|---|---|
| 10 | 94.7% | 91.2% | 12 | 18.3 |
| 25 | 93.8% | 89.4% | 14 | 42.6 |
| 50 | 92.6% | 87.1% | 17 | 89.1 |

### 4.5 Discussion

#### 4.5.1 Mechanism Analysis

FedNova-DP’'s performance improvements are due to two complementary mechanisms. On the first benefit, in FedAvg, if a single client has a very high percentage of extreme data distribution, their contribution to the gradient update is very large, so the direction of the contribution to the update comes from local data for that client, causing systematic error. First, normalized aggregation minimizes the impact that a large percentage of clients with extreme data distributions (producing very large gradient updates) have on the update direction of FedAvg. Whatever this bias may be due to, FedNova-DP absorbs it by normalizing each client's contribution by a corresponding local step count τi, assuming that both the amount of data on each individual client and the number of steps vary in a predictable manner. This is especially noticeable for the DP noise, where bigger gradient norms make larger noise term appear, and lead to a more biased sending direction of FedAvg’ s updates.

Second, adaptive clipping will minimize noise overhead when clients have a small gradient norm. Noise proportional to the fixed clipping threshold C is added irrespective of the presence of clipped

gradients, resulting in low signal to noise ratios with clipped gradients for clients whose gradients have small norms. FedNova-DP uses the median gradient norm to reduce the effect of low-frequency gradients the gradient is less than the gradient threshold (assisted by the specific calibration factor), resulting in a lower SNR in low gradient areas but a higher SNR in high gradient areas, thereby achieving the target mean SNR, especially for small and homogeneous patient populations.

### 4.5.2 Regulatory and Practical Implications

The classification of clinical decision support as a high-risk AI by the EU AI Act along with the FDA&#x2019s AI/ML-based framework for Software as a Medical Device (SaMD) more and more require formal mathematical guarantees of privacy for heuristic anonymization [1], [3]. In this section, FedNova-DP achieves clinical actionable results with (ε, δ)-DP (accuracy (ε, δ) = success with clinically acceptable accuracy ($\delta = 10^{-5}$), and with relatively strong privacy guarantees (ε = 0.5). Beyond ε = 2.0, the accuracy of FedNova-DP starts to decline, indicating a practical operating range of ε = 1.0–2.0, which guarantees good accuracy along with strong privacy. This is data-driven characterization which would provide organizations or institutions with some solid guidance on the decisions relating to privacy policies and deployment of clinical AI.

### 4.5.3 Limitations

Though common in FL research, the Dirichlet-based heterogeneity model may be inadequate to account for actual institution-based variation such as 'covariate shift' and changing coding practices over time as well as variations in protocols. Communication overhead analysis assumes perfect (no error, synchronous) communication scenario, in practice secure aggregation protocols are necessary that withstands an intermediate eavesdropper [22], [23], [24] and additional overhead are not addressed here. Lastly, the present study has been not extended to multi-label EHR classification, and extension to these tasks and other medical imaging applications such as survival analysis and time-series prediction are relevant topics for future research, in which gradient inversion attacks have already been well-documented [25].

## 5. CONCLUSION

This paper presented FedNova-DP: a privacy-preserving federated learning framework that use normalized weight aggregation, adaptive per-sample gradient clipping and calibrated Gaussian noise injection to ensure (ε, δ)-Differential Privacy. FedNova-DP substantially surpasses FedAvg-DP baseline with a 4.4 percentage-point increase of the performance across large-scale EHR classification task with rich privacy level (ε = 0.5) for heterogeneous client distributions and scalable sizes of the client federation, and is 31% faster when convergence is computed in communication rounds.

The results from privacy-utility analysis with different εs range {0.1, 10} indicate that the utility of FedNova-DP decreases as the privacy parameter ε increases with practical operating points of ε ≈ 1.0–2.0, achieving an accuracy of 91.7–93.2% while satisfying the healthcare AI governance regulations. Results from scalability tests verify that the benefits of the algorithm are sustainable up to 50 clients with graceful degradation of its performance from 10 clients. By examining failure modes in DP-FL, it is found that the normalized aggregation and adaptive clipping solve two different failure modes complimentarily and reinforcingly: first, gradient bias and second, excessive noise. The work demonstrates that in the realistic GA healthcare setting, managing privacy-utility tradeoffs in DP-augmented FL can also be quantitatively characterized and can be controlled without compromising the utility in practice. It offers concrete guidance on how to select the proper privacy budget, aggregation strategy, and scale the federation without degrading the utility of the FL system, as this would render it impractical to use. In the future, Fednova-DP will be developed asynchronously, hierarchically, into the personalization of institutions having patients with specific characteristics, and mathematically analyzed to withstand adaptive adversarial attacks against DP-FL gradient agents.

## Acknowledgments

The authors have no specific acknowledgments to make for this research.

## Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Author Contributions Statement

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Aruna Pavate | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

## Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Informed Consent

All participants were informed about the purpose of the study and their voluntary consent was obtained prior to data collection.

## Ethical Approval

Not applicable.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

[1] R. Miotto, 'Deep learning for healthcare: Review, opportunities and challenges', Brief. Bioinform, vol. 19, no. 6, pp. 1236-1246, 2018. doi.org/10.1093/bib/bbx044
[2] P. Rajpurkar et al., "AI in health and medicine," Nat. Med., vol. 28, no. 1, pp. 31-38, 2022. doi.org/10.1038/s41591-021-01614-0
[3] N. Rieke, 'The future of digital health with federated learning', vol. 3, no. 1, 2020. doi.org/10.1038/s41746-020-00323-1
[4] H. B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," AISTATS, 2017. doi.org/10.48550/arXiv.1602.05629
[5] L. Zhu et al., "Deep leakage from gradients," NeurIPS, vol. 32, 2019. doi.org/10.48550/arXiv.1906.08935
[6] J. Geiping et al., "Inverting gradients — How easy is it to break privacy in federated learning?" NeurIPS, vol. 33, 2020. doi.org/10.48550/arXiv.2003.14053
[7] C. Dwork and A. Roth, 'The algorithmic foundations of differential privacy', Found. Trends Theor. Comput. Sci, vol. 9, no. 3-4, pp. 211-407, 2014. doi.org/10.1561/0400000042
[8] X. Zhao, 'Local differential privacy for federated learning', ESORICS, 2022. doi.org/10.21203/rs.3.rs-1891162/v1
[9] T. Li et al., "Federated optimization in heterogeneous networks," MLSys, 2020. doi.org/10.48550/arXiv.1812.06127

[10] S. P. Karimireddy et al., "SCAFFOLD: Stochastic controlled averaging for federated learning," ICML, 2020. doi.org/10.48550/arXiv.1910.06378
[11] J. Wang et al., "Tackling the objective inconsistency problem in heterogeneous federated optimization," NeurIPS, vol. 33, 2020. doi.org/10.48550/arXiv.2007.07481
[12] M. Abadi et al., "Deep learning with differential privacy," ACM CCS, 2016. doi.org/10.1145/2976749.2978318
[13] R. C. Geyer et al., "Differentially private federated learning: A client level perspective," arXiv:1712.07557, 2017. doi.org/10.48550/arXiv.1712.07557
[14] N. Agarwal et al., "cpSGD: Communication-efficient and differentially-private distributed SGD," NeurIPS, 2018. doi.org/10.48550/arXiv.1805.10559
[15] I. Mironov, "Renyi differential privacy," IEEE CSF, 2017. doi.org/10.1109/CSF.2017.11
[16] N. Rieke, 'The future of digital health with federated learning', vol. 3, 2020. doi.org/10.1038/s41746-020-00323-1
[17] T. Li et al., "Federated learning for EHR-based sepsis prediction," JAMIA, vol. 28, 2021.
[18] P. Kairouz, 'Advances and open problems in federated learning', Found. Trends Mach. Learn, vol. 14, 2021. doi.org/10.1561/2200000083
[19] A. G. Roy et al., "BrainTorrent: A peer-to-peer environment for decentralized federated learning," arXiv:1905.06731, 2019. doi.org/10.48550/arXiv.1905.06731
[20] A. Hard et al., "Federated learning for mobile keyboard prediction," arXiv:1811.03604, 2018. doi.org/10.48550/arXiv.1811.03604
[21] V. Smith et al., "Federated multi-task learning," NeurIPS, 2017. https://doi.org/10.48550/arXiv.1705.10467
[22] W. Yang et al., "FFD: A federated learning based method for credit card fraud detection," BigData, 2019. doi.org/10.1007/978-3-030-23551-2_2
[23] Dayan, I., Roth, H. R., Zhong, A., *et al*., "Federated learning for predicting clinical outcomes in patients with COVID-19," Nature Medicine, vol. 27, no. 10, pp. 1735–1743, Oct. 2021, doi.org/10.1038/s41591-021-01506-3
[24] S. Truex et al., "A hybrid approach to privacy-preserving federated learning," ACM AIPri, 2019. doi.org/10.1145/3338501.3357370
[25] G. Kaissis et al., "Secure, privacy-preserving and federated machine learning in medical imaging," Nat. Mach. Intell., vol. 2, 2020. doi.org/10.1038/s42256-020-0186-1

**How to Cite:** Aruna Pavate. (2025). Privacy-preserving federated learning with differential privacy for healthcare AI: a convergence and utility analysis. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN), 5(2), 69–78. https://doi.org/10.55529/jaimlnn.52.69.78

**BIOGRAPHIE OF AUTHOR**

**Aruna Pavate**, is associated with the Department of Information Technology at Thakur College of Engineering and Technology, affiliated with University of Mumbai, India. Her academic interests include information technology, data analytics, artificial intelligence, machine learning, and emerging computing technologies. She has contributed to research focused on innovative technological solutions and interdisciplinary applications in modern computing systems. Through her academic and research activities, she actively supports advancements in technology-driven education and practical research initiatives aimed at addressing real-world challenges in the field of information technology. Email: arunaapavate@gmail.com