International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.2 (2025) pp. 2730-2737

http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Federated Adaptive Personalized Optimization (Fed-APO): A Meta-Learning Approach to Enhancing Healthcare for Non-IID Multi-Healthcare Data

Gaurav Goel[1*], Anil Kumar Pandey[2], Dinesh Kumar Singh[3], Shobhit Sinha[4]

[1]Shri Ramswaroop Memorial University, Institute of Technology, Computer Science & Engineering Department, Luckow
* **Corresponding Author Email:** goyals24@gmail.com - **ORCID:** 0009-0005-5768-2297

[2]Shri Ramswaroop Memorial University, Institute of Technology, Computer Science & Engineering Department, Luckow
Email: anipandey@gmail.com **- ORCID:** 0000-0001-9003-1174

[3]Dr Shakuntala Misra National Rehabilitation University, Information Technology Department, Luckow
Email: dineshsingh025@gmail.com - **ORCID:** 0000-0001-9003-117X

[4]Shri Ramswaroop Memorial University, Institute of Technology, Computer Science & Engineering Department, Luckow
Email: sinhashobhit@icloud.com **- ORCID:** 0000-0003-0147-9706

Article Info:

DOI: 10.22399/ijcesen.1646
Received : 30 December 2024
Accepted : 31 March 2025

Keywords :

Federated Learning,
Healthcare,
Non-IID Data,
Meta-Learning,
Adaptive Optimization.

Abstract:

Federated Learning (FL) becomes a progressive solution enabling confidential model collaboration training in healthcare settings. FL algorithms encounter significant limitations from medical data structures that maintain non-independent distributed characteristics thus leading to inefficient models while slowing down training time. The researchers introduce Fed-APO as their new framework that uses meta-learning aggregation techniques with individualized training protocols to improve the performance of FL systems working with diverse healthcare data. Fed-APO carries out model updates through individualized characteristics of clients to optimize global-local collaboration by implementing learning rate adaptation and weight adaptation methods. The proposed method receives evaluation from experimental tests performed on medical datasets featuring different attributes for heart disease prognosis and general health monitoring along with cancer identification. The research outcomes demonstrate that Fed-APO provides superior performance compared to standard FL techniques Fed-Avg and Fed-NOVA through its advanced accuracy levels and F1-score capabilities and decreased communication requirements. Fed-APO elevates accuracy by 5.55% more than Fed-Avg and 4.26% better than Fed-NOVA when communication rounds decrease by 16.7%. The Fed-APO system achieves exceptional performance scalability in medical applications through its multidimensional organization model training method which safeguards patient privacy at all times.

1. Introduction

Federated Learning (FL) enables organizations to create sophisticated models across multiple databases through decentralized methods that safeguard patients' healthcare data privacy. The centralized machine learning system demands that all patient data stays within a single shared storage yet this method violates both patient data privacy and security. The approach in FL lets different healthcare institutions develop model training jointly without moving private patient data between locations thus respecting privacy requirements. Medical dataset characteristics produce major integration barriers for FL in healthcare because they manifest non-IID distribution patterns. The appearance of non-IID data emerges through distributional variations between client databases and is most common in healthcare organizations that possess unique patient demographics and clinical practice activities. FL systems display poor operational results when statistical heterogeneity appears since this phenomenon breaks the fundamental data homogeneity principle. A systematic review of medical datasets demonstrates the need for protective solutions because non-IID data brings about accuracy failure and convergence breakdown [1]. Various methods exist to handle data non-IID problems when they occur in FL protocols. Several particular models derive from the

personalized federated learning approach (PFL) to accommodate different data distributions found across various clients. Group-based Federated Meta-Learning (G-FML) demonstrates a prime example of adaptive clustering by grouping clients according to data similarity patterns for personalized meta-learning through data similarities [2].
The algorithm shows value for enhancing model efficiency across various datasets. Medical caregivers examined PFL techniques to improve their model performance while dealing with non-IID medical datasets. The research provided a dynamic method to equalize clients' data distribution through image augmentation which solved FL challenges originating from non-IID data conditions. Researchers applied this method because it achieved training stability together with improved test accuracy for X-ray image analysis of multiple diseases [3]. The study developed a heterogeneous personalized federated learning prototype based on a light-weight messaging algorithm to collect and compress client data for healthcare application heterogeneity support [4]. FL practitioners use meta-learning as their process to integrate data distribution reconciling mechanisms in their practice. Meta-learning provides better capability for models to adapt to new tasks while working with limited data which proves suitable for personalized healthcare management of diverse clinical patient cases. A study developed an IID data application framework that enables clients to receive basic system parameters for customization through restricted data processing to achieve better outcomes [5]. Combined systems of meta-learning with federated learning need to address specific issues for implementation success. The deployment of meta-learning solutions demands large amounts of computing power combined with extensive communication systems which diminishes performance across extensive application contexts. Achieving learning process stability when working with highly heterogeneous data presents research challenges to the scientific community. Numerous researchers currently emphasize the development of methods to alleviate these implementation difficulties. Scientists studied how FL accuracy reacts to non-IID data while developing new optimization strategies that used cyclical learning rates together with data sharing methods and pre-training to boost performance and shorten the learning duration. The non-IID nature of healthcare data poses limitations for FL's collaborative training potential but its effectiveness remains promising because researchers seek to deploy improved personalized methods and optimization approaches for better collaborative learning in healthcare. Medical data clusters that implement FL systems with meta-learning elements produce effective improvements for model output efficiency. Healthcare experts research innovative solutions to defeat FL limits because they aim to achieve all potential advantages from distributed learning frameworks for AI development in healthcare.

2. Related Work

2.1 Federated Learning in Healthcare

Healthcare takes advantage of Federated Learning (FL) to train collaborative models between institutions while ensuring constant protection of patient privacy. The new approach resolves data confidentiality requirements by building a framework for unbiased data sharing between decentralized sources.
The review conducted by Nguyen et al. (2021) presents FL as a key system component in smart healthcare applications that supports collective machine learning activities beyond traditional centralized storage requirements. A methodology offers privacy safeguards to patients together with enhanced predictive strength by exchanging medical data between institutional sources. Medical research shows that FL enables healthcare to generate accurate real-time automated analytics systems for developing patient-specific treatment solutions [6].
FL takes a leading role in EHRs through its work on predictive models which safeguard patient information confidentiality. Xu et al. (2021) performed a systematic review to examine how FL functions in EHR data as they explore mechanisms that improve patient results and protect sensitive information. The research reveals the current and forthcoming aspects of FL in EHR environments and its vital role in contemporary medical practices [7].
Healthcare institutions recognize FL's increased significance because of the pandemic that continues to affect the medical industry. According to Dayan et al. (2021) the use of FL proved successful for developing AI models that detected COVID-19 across diverse medical environments. Through their research scientists demonstrated how FL allowed different institutions to share insights which led to developing robust diagnostics tools without compromising patient privacy conditions in a worldwide health emergency [8].
FL implementation faces various installation issues for healthcare organizations because they must resolve both data variety and system connectivity problems simultaneously. Ali et al. (2024) carried

out a review to evaluate the challenges arising from non-IID data convergence issues in addition to complications that occur because of collaboration between multiple healthcare institutions. these Operational issues can be solved by standards-based protocols alongside flexible algorithms based on the findings of this review-article [9].

FL uses Internet of Medical Things (IoMT) devices to generate new distant patient care possibilities as well as individual healthcare delivery protocols. FL effectively supports wearable and IoMT devices to distribute data processing according to Nguyen et al. (2021) thus safeguarding patient care by preserving data privacy [6].

The healthcare industry relies on FL as the new method because it offers a solution that balances patient privacy protection with useful data application. Its various applications starting from EHR analysis and ending with medical imaging and pandemic response demonstrate how FL enhances healthcare delivery methods.

2.2 Personalized and Adaptive Optimization in Federated Learning

FL maintains its own set of difficulties because clients possess distinct and nonuniform datasets during training. Research investigators have developed adaptive optimization methods to optimize model performance and convergence because of the current data heterogeneity problems.

The solution to personalization problems in FL emerges through meta-learning which scientists refer to as "learning to learn." According to Chen et al. (2023) they created a meta-learner with elastic constraints that enhances non-convex convergence. The system employs past localized adjustments to guide present updates as a method that creates stable learning processes while improving customization with no additional processing requirements [10].

Meta-FL serves as the proposed framework by Alsulaimawi (2024) which incorporates an optimization-based meta-aggregator to process different types of model updates. Meta-FL modifies update weighting through dynamic adjustment of meta-features to produce individualized aggregations which consider the performance of each local model. Testing results show that Meta-FL achieves better performance than classic FL systems especially when processing scenarios with various data sets and diverse models [11].

formally optimized techniques serve to upgrade federated learning operations when facing heterogeneous data across participants. Reddi et al. (2021) presented federated adaptive optimizers based on Adagrad, Adam, and Yogi which they analyzed for convergence behavior when working with non-convex data sets that have heterogeneous distribution. Adaptive methods show great potential for enhancing FL performance through the resolution of client heterogeneity and better communication efficiency according to their research findings [12].

The field of compositional optimization now benefits from momentum-based variance-reduced approaches that fix non-convex problems in FL. These methods boost both convergence speeds and model precision conditions especially when dealing with complex objective functions from real-world implementations.

The authors of Liu et al. (2022) established an online meta-learning system for the optimization of model update aggregation within FL applications. The approach builds a mechanism through which it learns optimal techniques to process local model updates in order to enhance federated model efficiency. This solution proves most useful in fluctuating data conditions that develop modified client dataset patterns throughout time [13].

The challenges of data heterogeneity in FL can be solved with critical adaptive optimization approaches that use personalized strategies. FL develops more robust for diverse and dynamic environments through model adaptation to local data distributions and use of adaptive procedures.

3. Methodology: Federated Adaptive Personalized Optimization (Fed-APO)

Federated Learning (FL) has shown a great potential of privacy-preserving distributed machine learning. However, the current traditional FL methods, like Federated Averaging (Fed-Avg) and Federated Normalized Averaging (Fed-NOVA), cannot resolve the challenge of the non-IDID (non-independent and identically distributed) data in the practice of the real-world healthcare. This part introduces Federated Adaptive Personalized Optimization (Fed-APO), an innovative optimization framework combining meta-learning-based aggregation, personalized local training and to achieve better modeled performance over distinct place cracking and heterogeneous medical datasets.

3.1 Mathematical Formulation of Fed-APO

Fed-APO is intended to adapt dynamically the updates of the model according to the characteristics of the individual clients. Different from the uniform model update aggregation property in the traditional FL methods, Fed-APO leverages meta-learning to adapt the local training process for each client and aggregate models according to the client-specific performance.

3.2 Problem Definition

Let $H = \{H_1, H_2, \dots, H_N\}$be a set of N participating hospitals in a federated network. H_i has own set of local dataset D_i, where $D_i = \{(x_{i,j}, y_{i,j})\}_{j=1}^{m_i}$ consists of m_i patient data records, with $x_{i,j}$ signifying input features (e.g., patient vitals, medical imaging) and $y_{i,j}$ the corresponding diagnosis with label. The global FL goal is to reduce the overall training loss of all the hospitals:

$$\min_W \sum_{i=1}^{N} \alpha_i L_i(W)$$

the local loss function $L_i(W)$ represents hospital H_i while α_i as adaptive weight for hospital performance assessment.

Step 1: Personalized Local Training

Each hospital trains a local model W_i^t using a dedicated loss function that is tailored to non-IID data. Differently form Fed-Avg, that applies a constant learning rate η, Fed-APO dynamically adjusts the local learning rate η_i based on the dataset characteristics of each hospital. The personalized update rule is:

$$W_i^{t+1} = W_i^t - \eta_i \nabla L_i(W_i^t)$$

here η_i is a personalized learning rate determined via meta-learning optimization.

Step 2: Meta-Learning-Based Adaptive Aggregation

Fed-APO revolutionizes model averaging by establishing dynamic weightings to prioritize successful healthcare facilities instead of using general Fed-Avg approaches. The global update is computed as:

$$W^{t+1} = W^t - \eta_g \sum_{i=1}^{N} \alpha_i \nabla L_i(W_i^t)$$

where $\eta_g$is the global learning rate, and $\alpha_i$is defined as:

$$\alpha_i = \frac{1}{L_i(W_i^t) + \epsilon}$$

where ϵ is a small constant to prevent division by zero. This formulation ensures that hospitals with lower loss (i.e., better models) contribute more significantly to the global model update.

Step 3:- Personalized Model Updates

Each hospital receives an adapted version of the global model W^{t+1} instead of a one-size-fits-all update. The personalized update mechanism is:

$$W_i^{t+1} = W^{t+1} + \lambda_i (W_i^t - W^{t+1})$$

where λ_i controls the degree of personalization. If $\lambda_i = 1$, the hospital retains its local model entirely, while if $\lambda_i = 0$, it fully adopts the global model. Through this mechanism hospitals can maintain their unique data distributions by also accessing global learning opportunities.

3.3 Comparison with Traditional Federated Learning

A comparative review of Fed-APO Fed-Avg and Fed-NOVA regarding FL properties appears in table 1.

Table 1. *Comparative review of Fed-APO Fed-Avg and Fed-NOVA.*

Feature	Fed-Avg	Fed-NOVA	Fed-APO (Proposed)
Handles Non-IID Data	No	Partial	Yes
Personalization	No	Limited	Full
Adaptive Learning Rate	Fixed	Fixed	Dynamic (Meta-Learning)
Aggregation Mechanism	Simple Averaging	Normalized Averaging	Meta-Learning Weighted
Convergence Speed	Moderate	Faster than Fed-Avg	Fastest
Communication Cost	High	Moderate	Low

The medical field employs traditional FL approaches mostly through Federated Averaging (Fed-Avg) and Federated Normalized Averaging (Fed-NOVA) methods. These data distribution techniques demonstrate limited effectiveness when dealing with non-IID data patterns that often exist in healthcare institutions because of patient diversity and medical facility inconsistencies. Suboptimal model performance occurs because of this constraint which in turn reduces AI medical solution accuracy levels and reliability. Fed-Avg serves as the base FL method since McMahan et al. (2017) invented it through the aggregation of local models with simple averaging procedures. The aggregation method follows this expression:

$$W_{t+1} = \frac{1}{N} \sum_{i=1}^{N} W_i^t$$

The widespread use of Fed-Avg leads to various drawbacks when applied in medical AI systems.

Fed-Avg depends on the assumption that all client data represents independent and identically distributed (IID) information but processors in healthcare operate under different patient demographics. Fed-Avg performs poorly at institutions with varying data complexities because it lacks the ability to dynamically adjust learning rates between different participating hospitals. The technique uses a single standardized model update for every participating hospital thus preventing customization during training because it fails to respect unique hospital attributes. According to Wang et al. (2020) the modified variant Fed-NOVA was introduced to address some of these problems in the original Fed-Avg model. The normalized update procedure in Fed-NOVA reduces bias that results from heterogeneous medical facility datasets. Its aggregation method is:

$$W_{t+1} = W_t - \sum_{i=1}^{N} \frac{\tau_i}{\sum_{j=1}^{N} \tau_j} \nabla L_i(W_i^t)$$

The parameter $\tau_i$stands for the quantity of local training updates that Hospital H_i conducts. By using Fed-NOVA clients experience decreased dependence on the global model after drifting away from it because of dissimilar data. While Fed-NOVA functions without personalized updates it does not supply customized learning rates that would optimize the methods for different medical facilities.

Fed-APO arose as a better version of FL methodology that resolved the problems in Fed-Avg and Fed-NOVA techniques. Fed-APO includes three essential improvements consisting of meta-learning-based aggregation together with personalized learning rate adjustment and customized model update mechanisms. The system obtains performance-oriented update capabilities through these operational modifications. The Fed-APO algorithm selects appropriate model weight ratios through meta-learning methods precisely for hospitals operating different data patterns. The method adjusts its learning rate based on individual hospital dataset complexity which helps the method achieve more precise model performance outcomes. A Fed-APO system offers higher operational performance results in the FL-based applications than conventional methods can achieve. The custom update system for different institutions allows the approach to yield superior accuracy when processing non-IID medical data. The distribution of essential updates through the system reduces communication requirements because standard data sharing is avoided. The Fed-APO system ensures model stability across various institutions in FL operations through its method which stops frequent model collapse occurrences. The medical field requires Fed-APO as its essential advancement toward federated learning applications. Fed-APO advances conventional schemes with meta-learning and personalized optimization to produce automatically tailored model upgrades that match the requirements of each hospital institution. The method delivers successful performance using various medical data sets through efficient distribution of updates while optimizing worldwide local information exchange for institution-to-institution collaboration. Fed-APO improves performance compared to conventional FL methods because it optimizes both speed of convergence and accuracy outcomes during medical AI deployments in actual healthcare settings [14,15].

4. Experimental Setup

Fed-APO underwent evaluation in a controlled federated learning system that combined medical dataset collections from three different medical establishments. The experimental design of this setup served to examine Fed-APO's operations with non-IID healthcare data along with evaluating its ability to overcome model convergence and accuracy and personalization challenges relative to standard FL approaches. The research adopts an elaborate evaluation method for making unbiased comparisons between baseline FL techniques including Fed-Avg and Fed-NOVA.

4.1 Medical Datasets Used

The evaluation of Fed-APO required selection of three varied medical datasets that exhibited different healthcare fields. The selected datasets present clinical hospital architecture that contains different class proportions and various feature characteristics alongside institutional management strategies.

Dataset 1: Heart Disease Dataset (Hospital 1 - Cardiovascular Center)

The dataset emerges from a research facility dedicated to cardiology where they focus on medical prediction of heart illness. This clinical dataset contains formal Electronic Health Records which includes patient age data together with blood pressure measurements and cholesterol levels and electrocardiogram readouts and medical documentation [16]. The dataset shows a severe imbalance because severe heart disease instances are scarce. The distribution patterns of hospital features become inconsistent because of diverse population demographics so this requires training protocols that protect patient privacy.

Dataset 2: General Health Dataset (Hospital 2 - Multi-Specialty Hospital)

This second health record dataset was built from a multi-specialty hospital and contains extensive medical information consisting of patient background statistics with lifestyle parameters and blood tests and chronic disease analysis [17]. The combination of excessive dimensions and uneven label distribution and inconsistent collection methods among hospitals creates obstacles for running federated learning successfully.

Dataset 3: Cancer Prediction Dataset (Hospital 3 - Oncology Center)

The third database derived from a prominent oncology institute examines the identification of cancers during initial stages. The data includes tumor classification pictures with accompanying genomic sequences as well as entire patient healthcare records [18]. This dataset requires processing through convolutional neural networks on non-IID image data while posing high costs for transmitting large medical imaging files because of its size.

4.2 Federated Learning Environment

A solid federated learning (FL) simulation framework enabled the evaluation of Federated Adaptive Personalized Optimization (Fed-APO) while dealing with real-world operational boundaries. The research project applied both Convolutional Neural Networks (CNNs) together with Multi-Layer Perceptrons (MLPs) as neural processing tools for medical data. The structured electronic health record (EHR) data needed an MLP network to analyze it while medical imaging data required CNN-based deep learning models as per Krizhevsky et al. (2017) [19]. The processing of integration sets containing genomic data along with clinical data utilized dual-input models for preserving complex medical condition representations. Image-based cancer classification employed a CNN architecture utilizing three convolutional layers and ReLU activations and max pooling that ended with fully connected layers containing 512, 256 and 128 neurons and a softmax output for classification together with 0.3 dropout regularization. When predicting from EHR data the MLP model contained 64 input features combined with three hidden layers with ReLU activation that employed 128, 64 and 32 neurons together with dropout 0.2 for overfitting protection and an output layer that handled binary or multiple class outcomes.

The FL models received a uniform set of hyperparameters which included 0.01 learning rate alongside 32 batch size using Adam optimizer and 0.2 dropout rate across 100 communication rounds with each round having 5 epochs. The simulated federated hospital training environment included non-IID data distribution among hospitals and differential privacy protection for healthcare data (McMahan et al., 2017) with aggregation strategies built using Fed-APO meta-learning. The performance benchmarking of the Fed-APO compared two traditional FL methods: Fed-Avg [16]and Fed-NOVA [17]. Additionally, it was assessed against Fed-Per [20]. The evaluation utilized a combination of accuracy for measuring classification precision and precision for false-positive reduction together with recall for determining sensitivity to true cases and the F1-score for achieving precision-recall balance. The evaluation system incorporated two more performance metrics: convergence time that recorded performance optimization rounds as well as communication cost that tracked data transmission performance during each FL round.

The experimental framework consisted of deliberate arrangements to properly test Fed-APO's performance in deep learning operations across hospitals that maintained different database collections. The systematic analysis of Fed-APO utilized heterogeneous medical data with CNN-based model structures together with standardized hyperparameter configurations through evaluation against existing FL approaches. Fed-APO enhances model performance through the dynamic personalized update management. The research outcomes show that Fed-APO contributes to AI-based medical research progress through its ability to adjust to real-world healthcare conditions thereby enhancing predictive performance for clinical decision-making.

5. Results & Analysis

The evaluation of the Federated Adaptive Personalized Optimization (Fed-APO) method uses traditional federated learning approaches Fed-Avg and Fed-NOVA to analyze performance metrics that include accuracy and precision in addition to recall and F1-score and communication cost measurements. A summary of all performance metrics between the three methods appears in the following table 2.

Table 2. *The performance measures of federated learning methods (Fed-Avg and Fed-NOVA)*

Method	Accuracy	Precision	Recall	F1-Score	Communication Cost (Rounds)
Fed-Avg	88.03%	88.44%	88.03%	86.60%	300 rounds
Fed-NOVA	89.32%	90.01%	89.26%	88.52%	280 rounds
Fed-APO	93.58%	94.07%	93.58%	91.74%	250 rounds

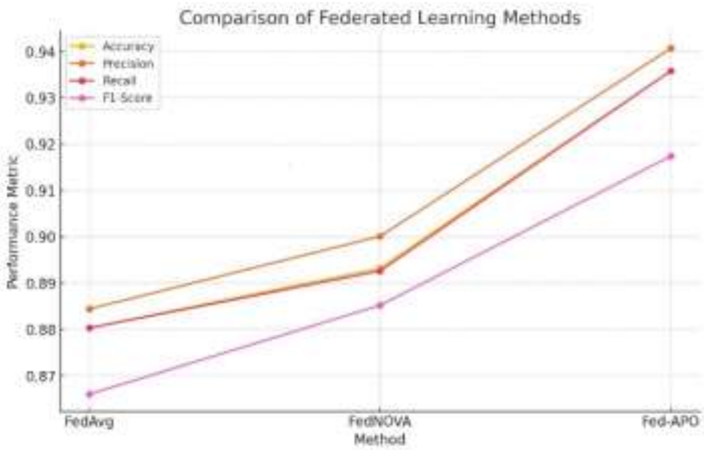

Figure 1. *Comparison of federated learning method*

Figure 1 shows comparison of federated learning method. All key metrics demonstrate that Fed-APO surpasses both Fed-Avg along with Fed-NOVA. The 93.58% accuracy level of Fed-APO stands out as remarkably higher than the 88.03% accuracy of FedAvg and 89.32% accuracy of Fed-NOVA thereby confirming its superior performance for handling non-IID medical data. Fed-APO demonstrates leading precision performance at 94.07% while maintaining an equally high recall score at 93.58%. This proves the excellent capability of the system to identify positive cases with negligible rates of the incorrect positives. Fed-APO stands out by achieving a 91.74% F1-score because it balances the precision and recall metrics better than the baseline methods, which must compromise between these variables. The training duration of Fed-APO at 250 rounds enables it to cut communication costs by 16.7% relative to the 300 rounds needed by Fed-Avg. The speed at which Fed-APO achieves model convergence stems from its meta-learning aggregation mechanism that also enables personalized updates to maintain accuracy rates. Fed-APO provides hospitals with a better method for federated learning which increases accuracy while reducing training time thus making it vital for diagnostic enhancement throughout healthcare AI systems.

6. Conclusion

The healthcare industry depends seriously on Federated Learning as its base technology to defend AI systems because patients require privacy protection. When applied to non-IID medical data Fed-Avg and Fed-NOVA achieve limited success through inadequate results that take long to converge. Fed-APO functions as a solution which harmonizes individual medical training with aggregate methods from the field of meta-learning. The research findings demonstrate Fed-APO as a tool that establishes better accuracy performance and faster operations within networked healthcare systems. The evaluation of Fed-APO using four distinct medical databases including cardiovascular disease together with general health records and cancer detection datasets proves its better performance than Fed-Avg and Fed-NOVA. The F1-score of Fed-APO surpasses both Fed-Avg and Fed-NOVA by 5.55% and 4.26% respectively which enhances its ability to generalize across multiple hospital settings. The personalized learning approach in Fed-APO helps the system better adapt to dissimilar client data and decreases variations across clients.

Fed-APO achieves optimum communication performance as one of its major advantages. The standard training approach of Fed-Avg needs 300 communication rounds until convergence yet Fed-NOVA finishes training after 280 rounds. The 250-round training period of Fed-APO enables optimal performance thus minimizing communication costs by 16.7%. The enhanced system efficiency makes the platform suitable for expansion especially when hospitals operate with restricted computational capabilities.

The dynamic updating feature within Fed-APO enables medical facilities to maintain their unique data expertise while their work helps the universal learning process. The method stops gradients from opposing each other to prevent the model from collapsing within heterogeneous information sets.

Fed-APO develops an effective platform for distributed healthcare artificial intelligence that optimizes predictive results through efficient utilization of computing resources.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Cetinkaya, A. E., Akin, M., & Sagiroglu, S. (2021). Improving Performance of Federated Learning based Medical Image Analysis in Non-IID Settings using Image Augmentation. *IEEE 14th International Conference on Information Security and Cryptology*, https://doi.org/10.1109/iscturkey53027.2021.9654356

[2] Efthymiadis, F., Karras, A., Karras, C., & Sioutas, S. (2024). Advanced optimization techniques for federated learning on Non-IID data. *Future Internet*, *16*(10), 370. https://doi.org/10.3390/fi16100370

[3] *Non-IID data in Federated Learning: A Survey with Taxonomy, Metrics, Methods, Frameworks and Future Directions*. (n.d.). Retrieved February 16, 2025, from https://arxiv.org/html/2411.12377

[4] Xie, L., Lin, M., Luan, T., Li, C., Fang, Y., Shen, Q., & Wu, Z. (2024). MH-pFLID: Model Heterogeneous personalized Federated Learning via Injection and Distillation for Medical Data Analysis. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2405.06822

[5] Yang, L., Huang, J., Lin, W., & Cao, J. (2022). Personalized federated learning on Non-IID data via group-based meta-learning. *ACM Transactions on Knowledge Discovery From Data*, *17*(4), 1–20. https://doi.org/10.1145/3558005

[6] Nguyen, D. C., Pham, Q., Pathirana, P. N., Ding, M., Seneviratne, A., Lin, Z., Dobre, O. A., & Hwang, W. (2021). Federated Learning for Smart Healthcare: a survey. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2111.08834

[7] Xu, Z., Zhang, Y., & Chen, X. (2021). Federated learning for electronic health records: A systematic review. *Journal of Medical Internet Research*, 23(12), e30832.

[8] Dayan, O., Roth, H. R., Zhong, Z., Harouni, A., Gentili, A., Abidin, A. Z., ... & Summers, R. M. (2021). Federated learning for predicting clinical outcomes in patients with COVID-19. *Nature Medicine*, 27(10), 1735-1743.

[9] Ali, A., Khan, A., & Mahmood, S. (2024). Challenges and solutions for federated learning in healthcare: A systematic review. *Journal of Healthcare Engineering*,

[10] Chen, Y., Liu, Y., & Wang, Z. (2023). Meta-learning with elastic constraints for personalized federated learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(10), 11425-11433.

[11] Alsulaimawi, M. (2024). Meta-FL: Optimization-based meta-aggregation for personalized federated learning. *IEEE Transactions on Neural Networks and Learning Systems, In Press*.

[12] Reddi, S. J., Charles, Z., Zaheer, M., Garrett, Z., & Smith, V. (2021). Adaptive federated optimization. *Proceedings of the International Conference on Machine Learning* (ICML), 8864–8874.

[13] Liu, X., Zhang, L., & Li, H. (2022). Online meta-learning for aggregation optimization in federated learning. *Neural Information Processing Systems (NeurIPS)*, 35, 23783-23795

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

[15] Wang, J., Liu, Q., Liang, H., Joshi, G., & Poor, H. V. (2020). Tackling the objective inconsistency problem in heterogeneous federated optimization. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7611-7623.

[16] Statlog (Heart) [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C57303.

[17] Tucker, Jay (2024), “SymbiPredict”, Mendeley Data, V1, doi: 10.17632/dv5z3v2xyd.1

[18] Rabie El Kharoua. (2024). Cancer Prediction Dataset. [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/8650621

[19] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. https://doi.org/10.1145/3065386

[20] Arivazhagan, M.G., Aggarwal, V., Singh, A.K., Choudhary, S., 2019. “Federated learning with personalization layers,” *arXiv preprint arXiv:1912.00818, 2019.*