Federated Learning Harnessed with Differential Privacy for Heart Disease Prediction: Enhancing Privacy and Accuracy

Wided Moulahi[1], Tarek Moulahi[2], Imen Jdey[3] and Salah Zidi[4]
[1]*IResCoMath Laboratory, National Engineering School of Gabes, Tunisia*
[2]*Department of Information Technology, College of Computer, Qassim University, Kingdom of Saudi Arabia*
[3]*ReGIM-Lab. REsearch Groups in Intelligent Machines (LR11ES48), Tunisia*
[4]*University of Haute Alsace, Colmar, 68000, France*

Keywords: Federated Learning, Differential Privacy, Health, Heart Disease, Privacy Preservation.

Abstract: The increasing digitization of healthcare raises the concerns surrounding the patients' privacy. Therefore, the integration of privacy preserving technologies has proven imperative to curb the negative repercussions tied to technology deployment in the medical sector and to provide trustworthy artificial intelligence healthcare applications. Two raising approaches are promoted to the forefront of research and gaining momentum in the realm of healthcare smart systems: Federated Learning and Differential Privacy. On one hand, Federated Learning (FL) enables collaborative model training across multiple institutions without exchanging raw data. Differential Privacy (DP), on the other hand, provides a formal framework for safeguarding data against potential privacy breaches. The application of these approaches in healthcare settings ensures the protection of sensitive patient informations. In this paper, we delve into the challenges posed by medical data to see how FL and DP can be tailored to suit these requirements. We aim to strike a balance between technology deployment in the medical field and privacy preservation. To this end, we developed a Multi-layer Perceptron (MLP) model to predict if a person is at risk to have heart diseases. The model, trained on different medical datasets for heart diseases, reached an accuracy of 99.57%. The same model was trained in FL framework. It achieved a FL averaged accuracy reaching 99.15%. In a third scenario, to enhance clients' privacy, we deployed a DP framework. The differentially_private MLP achieved an accuracy extending to 97.07% in centralized settings and averaged accuracy attaining 89.94% in FL settings, outperforming existing methods in heart diseases prediction.

1 INTRODUCTION

Machine Learning (ML) (Jordan and Mitchell, 2015) has the potential to transform healthcare industry by enhancing diagnosis, treatment, and patients' well-being. It is of a paramount importance in the medical field in diverse ways. One crucial area is diagnosis and disease prediction. In fact, ML algorithms are employed in drug discovery and development, where they assist in finding potential avenues for new treatments and improving existing medicines (Vamathevan et al., 2019; Brahmi et al., 2024). They also help to, early, identify diseases, especially those not easily detectable at an initial stage. Additionally, ML is used in medical imaging to aid pathologists in making more accurate diagnostic judgments. It streamlines routine tasks, enabling healthcare professionals to focus on essential aspects of patient care. It also assists in robotic surgeries, and identifies prescription errors (Kassahun et al., 2016).

However, these advancements come with hurdles, including the need for large, trustworthy datasets, the understanding of ML models, and ethical and regulatory issues. There are challenges that must be taken care of prior to the widespread integration of ML into clinical practice. The requirement for large, excellent datasets that truly portray the patient population is one of the main obstacles. Large amounts of patient data are crucial for the training and accuracy improvement of ML systems. Sensitive personal data, involving genetic, biometric, and medical history, is frequently included in this data. This brings up a number of issues with control, access, and data security (Hameed et al., 2021). There is a risk of re-identification and privacy breaches, given that individuals can potentially be identified through their data. According to a study

Moulahi, W., Moulahi, T., Jdey, I. and Zidi, S.
Federated Learning Harnessed with Differential Privacy for Heart Disease Prediction: Enhancing Privacy and Accuracy.
DOI: 10.5220/0013188900003890

In *Proceedings of the 17th International Conference on Agents and Artificial Intelligence (ICAART 2025) - Volume 3*, pages 845-852
ISBN: 978-989-758-737-5; ISSN: 2184-433X

conducted in the journal Nature Communications, approximately 99.98% of americans in an anonymized dataset could be re-identified (Rocher et al., 2019).

Furthermore, collaboration is needed between healthcare providers, researchers, and data scientists in order to collect and arrange data in a way that preserves data integrity and patient confidentiality. This is especially important in the medical field, where decisions made can have life-or-death consequences. To make well-informed judgments about patient care, physicians and other healthcare professionals ought to be able to trust and understand the decisions produced by ML models. Building trust in the technology depends on this transparency. Additionally, addressing regulatory and ethical considerations is crucial, such as algorithms' fairness, and protecting patient data. Despite these challenges, the potential benefits of ML in healthcare are significant, and with meticulous planning and collaboration, these hurdles can be overcome to improve patient care.

As the medical community continues to leverage data-driven approaches for improved diagnostics and treatment, the implementation of FL and DP emerges as a pivotal strategy to uphold the confidentiality and trustworthiness of patient informations. This work contributes to the ongoing discourse on privacy in healthcare by shedding light on the potential of cutting-edge technologies to revolutionize medical research and practice while steadfastly safeguarding individual privacy.

According to the World Health Organisation, the cardiovascular diseases are, globally, the leading cause of death. Each year, 17.3 millions of people die due to heart diseases[1].

Towards the aforementioned concerns, in this research paper:

1. We developed an efficient Multilayer Perceptron to predict whether a person is at risk to have heart diseases.
2. We enhanced the privacy of the approach by deploying it in a FL framework.
3. We optimized the privacy preservation by deploying DP approach.

The remaining parts of this paper are organized as follows: in section 2, we summarize the related work. In section 3, we outline the problem statement. Section 4 introduces the proposed contribution. The results of our approach are presented and discussed in section 5, and we conclude with section 6.

2 RELATED WORK

Several privacy preserving smart frameworks were developed to tackle medical problems while maintaining a balance between ML efficiency and privacy preservation in smart healthcare systems. Table 1 summarizes some of these frameworks.

(Wang et al., 2024) proposed a Differentially Private Federated Transfer Learning Framework using MLP for stress detection. This approach combines DP with FL. Its accuracy was reported to be 53%. One limitation of this approach could be the relatively lower accuracy of 53% achieved in stress detection.

(Savić et al., 2023) proposed ML techniques to predict Quality of life (QoL) indicators for cancer patients using centralized and FL scenarios for model training on ORB and BcBase datasets. Different ML models were used for classification and regression tasks. Centralized and federated models show comparable predictive power for QoL. Optimal privacy values for regressors show steady mean absolute error (MAE) value decrease.

The work of (Babu Nampalle et al., 2023) integrated DP into FL for medical image classification. The developed model was based on noise calibration, adaptive privacy budget strategy, and privacy-utility trade-off analysis. It used the MobileNetV2 pre_trained model on HAM10000 Skin Image Dataset, four discrete datasets from the Cancer Imaging Archive, PH2 and Memorial Sloan Kettering datasets for skin images. The proposed framework achieved the following results of accuracy: 90.68%, 88.21% and 84.64%.

(Letafati and Otoum, 2023) proposed a distributed DP mechanism for metaverse healthcare using 'mix-up' noise. The model was evaluated on Breast Cancer Wisconsin Dataset addressing privacy-utility trade-off and diagnosis accuracy. The authors conducted the research over different numbers of clients for different levels of privacy and compared private scheme with non-private centralized setup for diagnosis accuracy.

(Liu et al., 2024) proposed a Record-Level Personalized DP (rPDP-FL) FL framework. It was applied on two datasets namely Fed-Heart-Disease and MNIST. For Fed-Heart-Disease, the accuracy reported varies between 77.17% and 81.89%. On the MNIST dataset, the accuracy varies between 84.11% and 94.77%.

[1] https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

Table 1: Privacy preserving frameworks applied in the medical field.

Ref	Method used	Results	Limitation
(Wang et al., 2024)	A differentially private federated transfer learning framework using MLP for stress detection	Accuracy: 53% ROC Curve: NON-DP : Area=0.59 Epsilon=1 : Area=0.56 Epsilon=0.5 : Area=0.54	Relatively low accuracy of 53% achieved in stress detection.
(Savić et al., 2023)	Application of ML techniques to predict the quality of life features for patients diagnosed with cancer	The accuracy varies between 51.6% and 71.4% among the different datasets and the different ML models. The MAE values vary between 5.1055 and 6.7250 for different values of DP and different regressors	The performances of the models could be enhanced.
(Babu Nampalle et al., 2023)	FL framework with integrated DP for medical image classification using MobileNetV2 architecture	Accuracy: Baseline: 90.68% , FL: 88.21% , DP_FL: 84.64%	The system could be improved to enhance performance, privacy and security.
(Letafati and Otoum, 2023)	Distributed Differential Privacy for the metaverse healthcare systems for breast cancer diagnosis	$\varepsilon = 20$: 62% of accuracy. $\varepsilon = 60$: 90% of accuracy.	The system performance is enhanced on behalf the privacy. Find a trade-off between accuracy and privacy budget.
(Liu et al., 2024)	Federated Learning framework based on record-level personalized Differential Privacy(referred to as rPDP-FL) applied on two datasets.	Fed-Heart-Disease: Accuracy: from 77.17% to 81.89% . MNIST: Accuracy: from 84.11% to 94.77%	The performances of the FL could be enhanced.

3 PROBLEM STATEMENT

Two key approaches are the pillars of our research: FL and DP.

3.1 Federated Learning

FL is a game-changing concept in ML. Unlike traditional methods that rely on centralized data, it allows multiple devices to collaborate on training models. This approach is particularly valuable in scientific domains because it protects privacy by keeping data on local devices instead of being shared (Moulahi et al., 2023). Depending on clients' interaction with the process, FL can be conducted in three modalities: synchronous FL, asynchronous FL and semi-synchronous FL.

In FL, the participating devices might have heterogeneous computation potentials. The synchronous FL does not consider how heterogeneous these devices are. The lowest device determines the speed of the process. The devices with the highest computational resources remain idle until the other devices achieve their local training (Feng et al., 2021; Stripelis et al., 2022).

In contrast, the asynchronous FL does not synchronise the communication between the different participating devices. Once it achieves its training, each device uploads its local updates and downloads the new updated model without waiting the other devices (Feng et al., 2021; Stripelis et al., 2022).

Semi-synchronous FL combines features of synchronous and asynchronous methods. It allows devices to synchronize periodically with a central server or each other. Semi-synchronous FL balances speed and synchronization in FL (Feng et al., 2021; Stripelis et al., 2022).

3.1.1 FL Security and Privacy Issues

FL faces challenges which call its efficiency into question. One of these threats are Data poisoning attacks which aim to degrade the model performances. It consists to inject carefully crafted samples in the dataset to mislead the model behaviour. These samples could be injected in the training dataset or in the testing dataset (Sun et al., 2022). The Model poisoning attacks aim to modify the model parameters and learning rule. These attacks could be injected by a malicious client or a malicious server (Sun et al.,

2022). Inference attacks are adversarial algorithms that trace back the samples in the training dataset. They aim to divulge the private informations of participants (Jdey, 2022). Byzantine attacks aim at degrading the global model convergence. Malicious clients falsify the data or the model updates so that the model converges slowly (Prakash and Avestimehr, 2020). Free-riding attacks intend to obtain the final global model without participating in the training process (Bouacida and Mohapatra, 2021).

3.2 Differential Privacy

DP is a way to keep data private in data analysis and ML. It consists of adding an amount of noise to the data or the results of algorithms (Dwork, 2006). Technically, DP is based on the idea of "neighboring datasets." Two datasets are considered neighbors if they only differ by one entity.

DP is a privacy-preserving method that quantifies how much privacy is lost. It's measured by a parameter called epsilon (ε), which represents the maximum privacy loss allowed (Lee and Clifton, 2011). A larger ε means less privacy but more useful data. DP helps strike a balance between using data to learn meaningful insights and protecting the privacy of individuals represented in the dataset (Dwork et al., 2014).

4 PROPOSED CONTRIBUTIONS

4.1 Contributions Description

Our approach comprises three distinct processes aimed at predicting heart diseases while preserving privacy.

The first process consists of developing an efficient MLP model on different heart disease datasets to generate predictive insights regarding the likelihood of heart diseases occurrence in individuals. Two tasks are targeted : binary classification and multiclass classification.

Following the initial MLP modeling, we embark on the second process, which employs FL approach. We tend to preserve clients' privacy through collaborative learning, each data source contributes knowledge to the model without exposing individual records. During this process, several communication rounds take place. A single cycle consists of clients carrying out local computations on their data and then transmitting the updates to the central server for aggregation.

In the final process, to optimize privacy preservation, we integrate DP mechanism. We apply DP on the MLP model in centralized settings. This involves adding noise to the model parameters to protect them against attacks and preserve the data privacy. The differentially private MLP model is then used for making predictions, while still preserving data stakeholders' privacy.

The DP technique is, then, extended to FL where the MLP model is trained across the different clients. We ensure, thereby, protecting the confidentiality of sensitive information while allowing effective model training and prediction. We are applying a local DP where noise is added locally on each client model parameters before sending them to the central server for aggregation.

The results of our approach are discussed comprehensively, considering both predictive performance and privacy preservation. We analyze the accuracy and robustness of the MLP model trained on different heart disease datasets, evaluating its efficacy in identifying individuals at risk of heart diseases. Additionally, we assess the impact of FL and DP on model performance and privacy preservation, highlighting any trade-offs and benefits observed. Figure 1 presents the approach our research adopted.

4.2 Suggested Scenarios

4.2.1 Process_1: Centralized Learning

A main goal of our research is to ensure privacy preservation. Therefore, the datasets we used are medical datasets for heart diseases:
Heart_1:www.kaggle.com/datasets/johnsmith88/heart-disease-dataset.
Heart_2:https://www.kaggle.com/datasets/sid321axn/heart-statlog-cleveland-hungary-final.
Heart_3:https://archive.ics.uci.edu/dataset/193/cardiotocography.

These datasets are used to train a supervised MLP model (Popescu et al., 2009). Its parameters are described in Table 2.

4.2.2 Process_2: Federated Learning

Our collaborative model involves different numbers of clients. In a scenario, on the first dataset (Heart_1), five clients are involved. In an other scenario (Heart_2), we involved two clients, and in the last scenario (Heart_3), three clients were involved. The training process goes through 10 rounds. For averaging updates coming from clients, we used FedAvg as an aggregation algorithm (Issa et al., 2023). All clients participate in each communication round simultaneously. This synchronous approach ensures

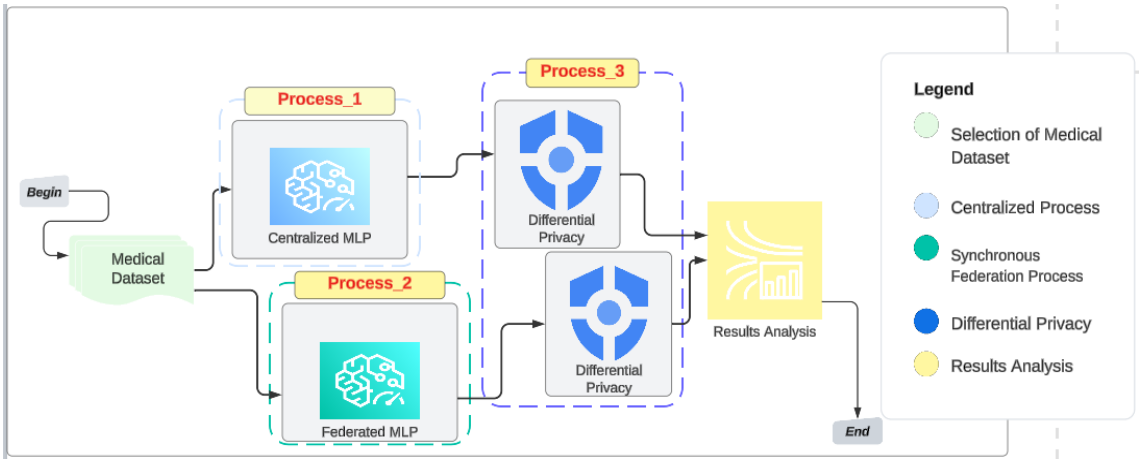

Figure 1: The proposed Approach which comprises three processes: centralized Learning, FL and DP integration.

Table 2: MLP hyper-parameters values.

Hyper-parameter	Heart_1	Heart_2	Heart_13
Input layers	13	11	20
Hidden layers	9	9	9
Optimizer	Adam	Adam	Adam
Loss Function	Binary_crossentropy	Binary_crossentropy	Sparse_categorical_crossentropy
Epochs	20	20	20
Activation function	Relu, Sigmoid	Relu, Sigmoid	Relu, Softmax
Classes	2	2	3
Task	Binary classification	Binary classification	Multi-class classification

that all clients are updated with the latest global model before proceeding to the next round.

4.2.3 Process_3: Differential Privacy Integration

The DP integration process takes place using different DP parameters (Table 3). We used the PyTorch Opacus library for DP deployment. The performances of the differentially private MLP model in both centralized and federated settings are then analysed to measure the trade-off between model accuracy and the level of privacy achieved. The results provide insights into the effectiveness of the proposed DP technique in preserving privacy while maintaining model performance. “Small” ε values (canonically, $\varepsilon \leq 1$) tend to exhibit great privacy guarantees but often severely impact performance. ”Medium” and ”large” ε values (canonically, $\varepsilon \in [1, 10]$ and $\varepsilon \geq 10$, respectively) provide more relaxed privacy guarantees, but increase utility (Lee and Clifton, 2011)

5 RESULTS AND DISCUSSION

5.1 Results

To evaluate the performances of our approach, we used the following metrics: accuracy, precision, recall, F1 score, F-beta score and Receiver Operation Characteristic curve (ROC Curve) (Awad and Hassanien, 2014).

After implementing the previously proposed approach, it achieved the following results: Table 4 shows the performances of the MLP model trained on the selected datasets in centralized settings as well as in federated settings. Figure 2 depicts the ROC Curves of the three datasets in centralized settings. Figure 3 depicts the ROC Curves of the three datasets in centralized settings after applying DP using different DP parameters values. Table 5 highlights the model performance after applying DP in centralized and federated settings on the three datasets. It presents the model accuracy and privacy budgets spent for different parameters.

Table 3: DP parameters.

Parameter	**Description**
Noise_multiplier (σ)	Amount of noise added to gradients
Delta (δ)	Desired upper bound on the probability of information leakage
Alpha (α)	Level of privacy protection
Epsilon (ε)	Privacy budget

Table 4: MLP performances in centralized settings and federated settings on different datasets.

Dataset	Heart_1		Heart_2		Heart_3	
Metric	Centralized	FL	Centralized	FL	Centralized	FL
Accuracy	99.51%	98.86%	99.57%	99.15%	98.82%	97.08%
Precision	98.98%	99.23%	100%	100%	99.16%	97.29%
Recall	100%	98.45%	99.20%	98.35%	99.12%	98.13%
F1 score	99.49%	98.81%	99.59%	99.17%	99.13%	98.01%
F Beta score	99.19%	99.05%	99.83%	99.66%	99.15%	98.75%

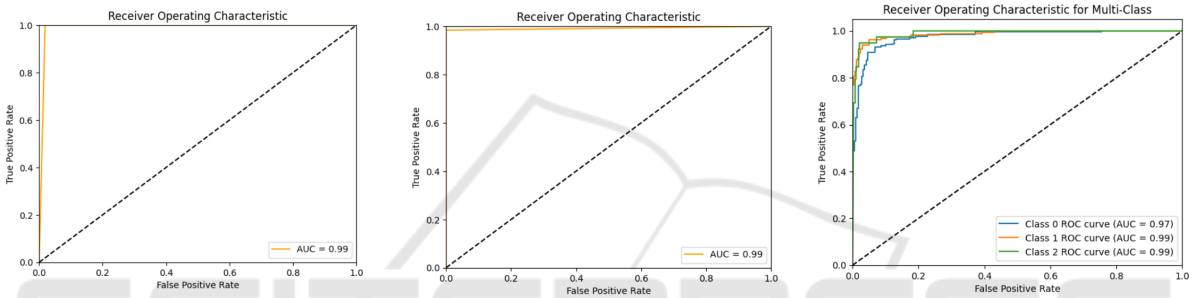

Figure 2: ROC curves of, respectively, Heart_1, Heart_2, Heart_3.

Table 5: Model performance for different datasets and different privacy budgets with $\delta = 1e-5$ and sample rate (q)=0.01.

Parameters	$\sigma = 1.3$ $\alpha = 5$ $\varepsilon = 2.25$		$\sigma = 1.3$ $\alpha = 10$ $\varepsilon = 0.91$	
Dataset	DP_Centralized	DP_FL	DP_Centralized	DP_FL
Heart_1	92.68%	70.12%	97.07%	75.95%
Heart_2	86.97%	89.04%	86.97%	88.94%
Heart_3	95.06%	89.94%	94.12%	89.71%

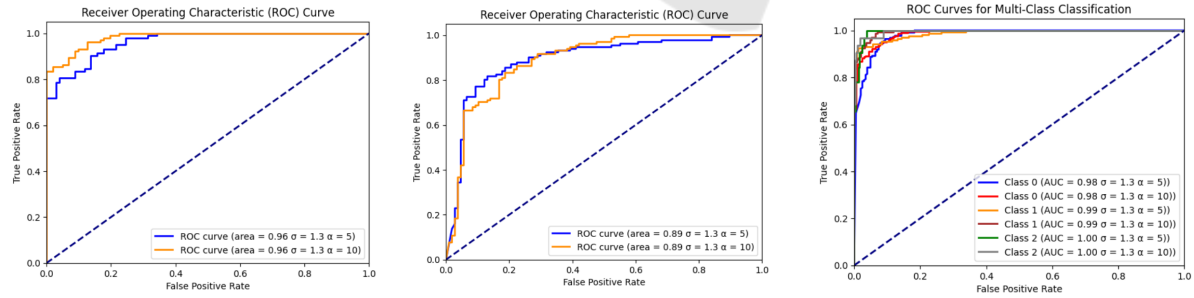

Figure 3: ROC curves of, respectively, Heart_1, Heart_2, Heart_3 in centralized settings after applying DP mechanism with different values of DP parameters.

5.2 Discussion

Building upon the results presented in the previous section, the developed MLP model shows efficient results in terms of accuracy, precision, recall, F1 score and F-beta score and ROC Curve. The high values of these performance metrics indicate that the MLP model achieved reliable results in predicting heart disease risk with a high degree of accuracy and reliability. Along with, the accuracy of 99.57%, which measures the model's ability to make correct predictions out of the total predictions, signifies that our MLP has

excellent discriminatory power.

Furthermore, in federated settings, deploying the MLP model helps preserving clients' privacy without sacrificing the model performances. Despite the distributed nature of data, the FL still achieves competitive accuracy, showcasing the effectiveness of collaborative model training and its ability to maintain high performance levels across repeated interactions and updates during the FL process. Additionally, figure 4 exhibits the closely related performances between the centralized learning and FL while providing more privacy to data stakeholders.

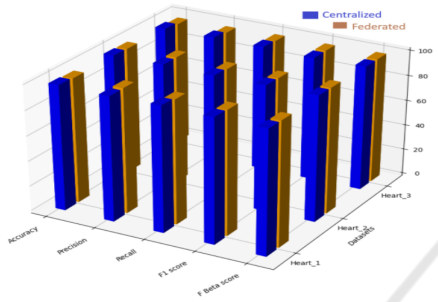

Figure 4: MLP performances in centralized settings against MLP performances in federated settings on the three datasets.

The differences between the performances of centralized DP_free MLP and the differentially private MLP performances are reimbursed by the more security and privacy added to the individuals data. Carefully tuned DP parameters help, considerably, achieving the desired level of privacy while maintaining reliable levels of accuracy. With σ=1.3, α=10, the consumed privacy budget (ε=0.91) is still in an acceptable range, for a satisfying accuracy reaching 97.07%.

Highlighting the difference between the accuracy of the DP_Free MLP and the accuracy of differentially private MLP trained using different parameters in federated settings, the performances of the model are still accurate despite the fact that two privacy preserving mechanisms are deployed. The slight decrease in model performances is indemnified by a strong privacy guarantee due to the use of DP harnessed with FL. With σ=1.3, α=10, the consumed privacy budget (ε=0.91) is still in an acceptable range for an accuracy of 89.71% .

Our approach, applied on different heart diseases datasets, proves its scalability and shows very satisfying results outperforming existing results. Extended to a multi-classification task, it achieves reliable and accurate results.

Compared to the work of (Liu et al., 2024), our work provides a three-dimensional approach which involves centralized, federated and differentially_private models. That allows examining the impact of each technique on the learning process. Besides, the FL performances of our approach exceed those of the aforementioned work. Our model achieves an average accuracy of 99.15% against an average accuracy less than that of (Liu et al., 2024). Actually, the concerned paper does not provide concrete values of the results but rather plots them , which make the comparison, a bit, confusing. For the federated learning, the plotted result seems less than 99%. Furthermore, the result of our differentially_private model exceed the results of his approach. Our proposed approach achieved more accurate results for less privacy budget. Our DP_MLP accuracy in centralized setting reaches 97.07% for a privacy budget $\varepsilon = 0.91$ against an accuracy of 81.34% for $\varepsilon \geq 1$.

6 CONCLUSION

Our research has demonstrated the successful deployment of a MLP model to predict the likelihood of individuals to have heart disease. Through the integration of FL and the incorporation of DP techniques, individual data contributions remain indistinguishable. We have not only achieved promising predictive performance but also preserved the privacy of sensitive medical data. Moreover, by integrating DP into the training process, we ensured that the model's predictions did not compromise the confidentiality of sensitive informations. DP mechanism comes to protect the model parameters against the attacks that may lead to data re-identification.

Our experimental results indicate that the deployed MLP model achieved satisfying levels of accuracy. That ensures its scalability and generalization across diverse datasets, underscoring its potential for real-world application in healthcare field. Furthermore, the successful implementation of privacy-preserving techniques highlights the feasibility of leveraging advanced ML methods while upholding strict privacy standards.

Overall, our findings contribute to the growing body of research on privacy-preserving techniques in ML and underscore the potential of FL combined with DP in healthcare applications. Moving forward, further exploration and refinement of our proposed approach hold promise for advancing predictive efficiency and privacy preservation in medical data analysis by applying other DP perturbation mechanisms in global or distributed DP architectures harnessed with other aggregation algorithms.

REFERENCES

Awad, A. I. and Hassanien, A. E. (2014). Impact of some biometric modalities on forensic science. *Computational intelligence in digital forensics: Forensic investigation and applications*, pages 47–62.

Babu Nampalle, K., Singh, P., Vivek Narayan, U., and Raman, B. (2023). Vision through the veil: Differential privacy in federated learning for medical image classification. *arXiv e-prints*, pages arXiv–2306.

Bouacida, N. and Mohapatra, P. (2021). Vulnerabilities in federated learning. *IEEE Access*, 9:63229–63249.

Brahmi, W., Jdey, I., and Drira, F. (2024). Exploring the role of convolutional neural networks (cnn) in dental radiography segmentation: A comprehensive systematic literature review. *Engineering Applications of Artificial Intelligence*, 133:108510.

Dwork, C. (2006). Differential privacy. In *International colloquium on automata, languages, and programming*, pages 1–12. Springer.

Dwork, C., Roth, A., et al. (2014). The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.

Feng, L., Zhao, Y., Guo, S., Qiu, X., Li, W., and Yu, P. (2021). Blockchain-based asynchronous federated learning for internet of things. *IEEE Transactions on Computers*, 99(1):1–9.

Hameed, S. S., Hassan, W. H., Latiff, L. A., and Ghabban, F. (2021). A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. *PeerJ Computer Science*, 7:e414.

Issa, W., Moustafa, N., Turnbull, B., Sohrabi, N., and Tari, Z. (2023). Blockchain-based federated learning for securing internet of things: A comprehensive survey. *ACM Computing Surveys*, 55(9):1–43.

Jdey, I. (2022). Trusted smart irrigation system based on fuzzy iot and blockchain. In *International Conference on Service-Oriented Computing*, pages 154–165. Springer.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260.

Kassahun, Y., Yu, B., Tibebu, A. T., Stoyanov, D., Giannarou, S., Metzen, J. H., and Vander Poorten, E. (2016). Surgical robotics beyond enhanced dexterity instrumentation: a survey of machine learning techniques and their role in intelligent and autonomous surgical actions. *International journal of computer assisted radiology and surgery*, 11:553–568.

Lee, J. and Clifton, C. (2011). How much is enough? choosing ε for differential privacy. In *Information Security: 14th International Conference, ISC 2011, Xi'an, China, October 26-29, 2011. Proceedings 14*, pages 325–340. Springer.

Letafati, M. and Otoum, S. (2023). Digital healthcare in the metaverse: Insights into privacy and security. *IEEE Consumer Electronics Magazine*.

Liu, J., Lou, J., Xiong, L., Liu, J., and Meng, X. (2024). Cross-silo federated learning with record-level personalized differential privacy. In *Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security*, pages 303–317.

Moulahi, W., Jdey, I., Moulahi, T., Alawida, M., and Alabdulatif, A. (2023). A blockchain-based federated learning mechanism for privacy preservation of healthcare iot data. *Computers in Biology and Medicine*, 167:107630.

Popescu, M.-C., Balas, V. E., Perescu-Popescu, L., and Mastorakis, N. (2009). Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, 8(7):579–588.

Prakash, S. and Avestimehr, A. S. (2020). Mitigating byzantine attacks in federated learning. *arXiv preprint arXiv:2010.07541*.

Rocher, L., Hendrickx, J. M., and De Montjoye, Y.-A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature communications*, 10(1):1–9.

Savić, M., Kurbalija, V., Ilić, M., Ivanović, M., Jakovetić, D., Valachis, A., Autexier, S., Rust, J., and Kosmidis, T. (2023). The application of machine learning techniques in prediction of quality of life features for cancer patients. *Computer Science and Information Systems*, 20(1):381–404.

Stripelis, D., Thompson, P. M., and Ambite, J. L. (2022). Semi-synchronous federated learning for energy-efficient training and accelerated convergence in cross-silo settings. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(5):1–29.

Sun, Y., Ochiai, H., and Sakuma, J. (2022). Semi-targeted model poisoning attack on federated learning via backward error analysis. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., Li, B., Madabhushi, A., Shah, P., Spitzer, M., et al. (2019). Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery*, 18(6):463–477.

Wang, Z., Yang, Z., Azimi, I., and Rahmani, A. M. (2024). Differential private federated transfer learning for mental health monitoring in everyday settings: A case study on stress detection. *arXiv preprint arXiv:2402.10862*.