Efficiency and Privacy in Deep Meta-Learning

Roushan Sherief (*Computer Science and Engineering*), Eng. Mayar Al-fares (Digital Media Engineering) and **Prof. Mohammed Salem** (*Digital Media Engineering*)

The German University in Cairo

roushan.abdelmaksoud@student.guc.edu.eg

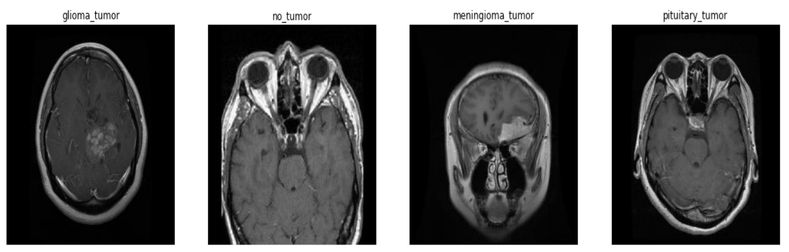
Brain tumors are one of the most prevalent neurological illnesses. Early discovery of a brain tumor aids radiologists in making an accurate prognosis and increasing the chances of long-term survival, although it is still a difficult process due to the tumor's changing appearance, location, form, and size. The most significant disadvantage of Deep Learning (DL) is that it necessitates a large amount of labeled data. Whereas medical datasets are typically modest. Thus, a DL model is not able to quickly and efficiently learn to recognize brain tumor categories. Medical data is particularly confidential, so these regulatory protocols are critical. From a privacy standpoint, the data that a centralized DL model may be able to collect and the potential uses for that data are causing increasing concern.

Literature Review

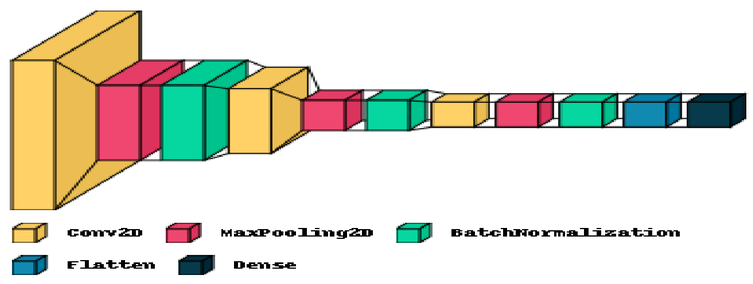
Kaissis et al. presented Privacy-preserving Medical Image Analysis (PriMIA) [1]. PriMIA is a commercial, free software platform that enables differentially private, securely aggregated Federated Learning (FL) and secured inferences on clinical data. The case study of PriMIA was presented on the pediatric pneumonia dataset [2] by training an 11.1 million parameter ResNet18 Convolutional Neural Network (CNN) [3]. This framework is compatible with a wide range of medical imaging data formats, is simple to configure, and improves FL training functionality. In situations when data cannot be uploaded to preserve clients' privacy, FL is a very promising solution for distributed Machine Learning (ML). As a result, FL is highly suited for real-world applications like analyzing sensitive healthcare data. McMahan et al. [4] used the FedAvg algorithm to conduct an early examination of FL. FedAvg works by having a coordinating server initialize a model before distributing it to clients. Clients run numerous epochs of Stochastic Gradient Descent (SGD) on their local datasets using the FedAvg [4] algorithm. Clients send their models to the server, which then averages them to create a new global model. The issue of lack of client adaptivity was addressed by Reddi et al. [5] and a basic structure for introducing adaptivity into FL is presented. They suggested a universal optimization framework (FedAdam) in which (1) clients undertake numerous epochs of training using a client optimizer to minimize a loss of their local data, and (2) the server updates its global model by averaging the clients' model modifications using a gradient-based server optimizer. Fallah et al. [6] investigated a personalized variant of FL (Per-FedAvg). The goal was to create an initial shared model that current or new users can quickly modify to their local datasets. This could be done by completing one or a few gradient descent steps on their own data. This technique retained all of the merits of FL architecture while also resulting in a more personalized model for each user due to its structure. They explained how the Model Agnostic Meta-Learning (MAML) framework can be used to investigate this issue. Whereas, the researchers contributed to personalizing FL using Moreu Envelopes [7] as loss functions and Adam optimizers [8] on the client side.

Methodology

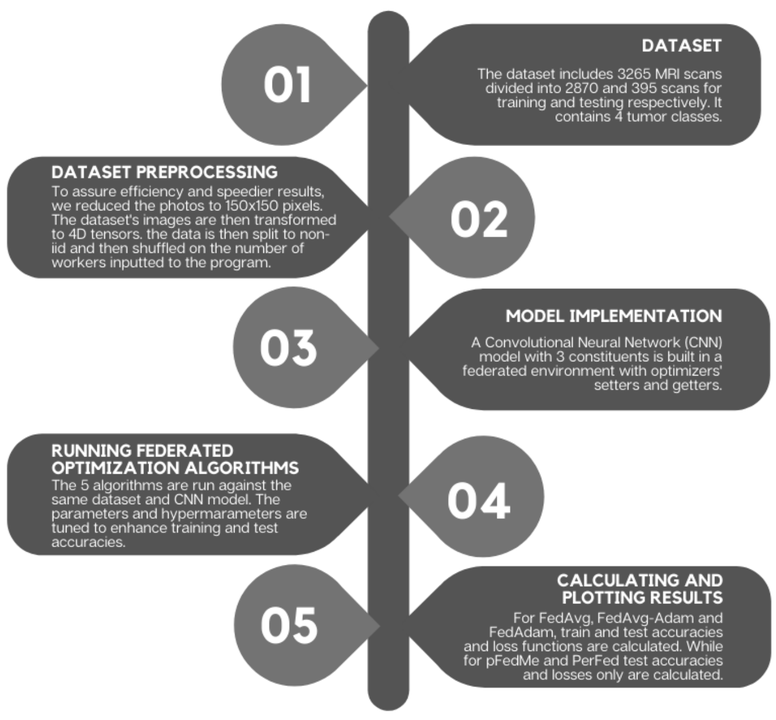
The Brain Tumor MRI dataset [9] contains MRI data. The images are split into Training and Testing folders. Each folder has four subfolders. These folders have MRIs of respective tumor classes. The four classes are Glioma Tumor, No Tumor, Meningioma Tumor, and Pituitary Tumor.



The model is a CNN model. It consists of 3 constituents and every constituent is made up of 4 layers as shown.

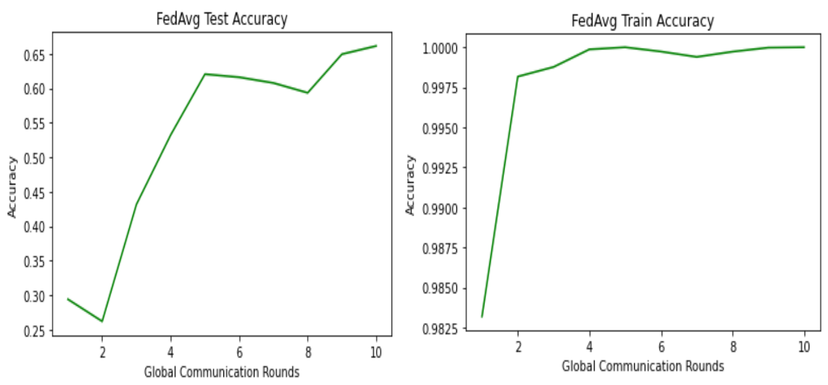


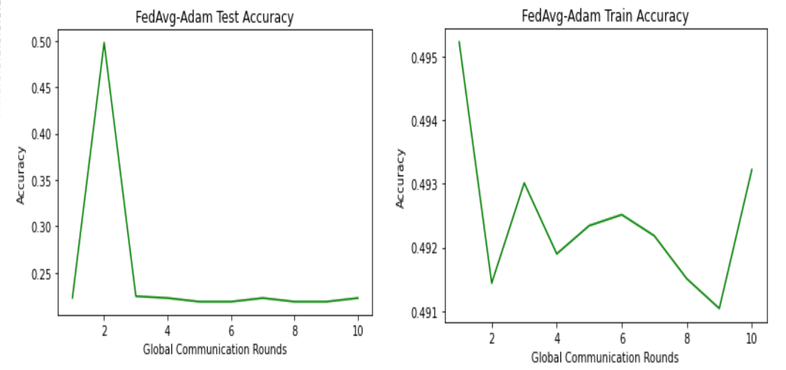
Our aim here in this research is to take [4][5][6][7][8]'s work which is based on optimizing and personalizing FL to the next step. We ran the 5 algorithms FedAvg, FedAvg-Adam, FedAdam, Per-FedAvg, and pFedMe) on the MRI dataset [9] and our CNN model.

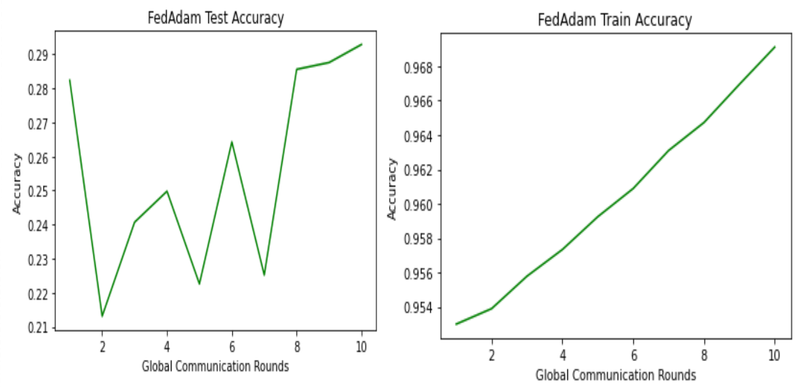


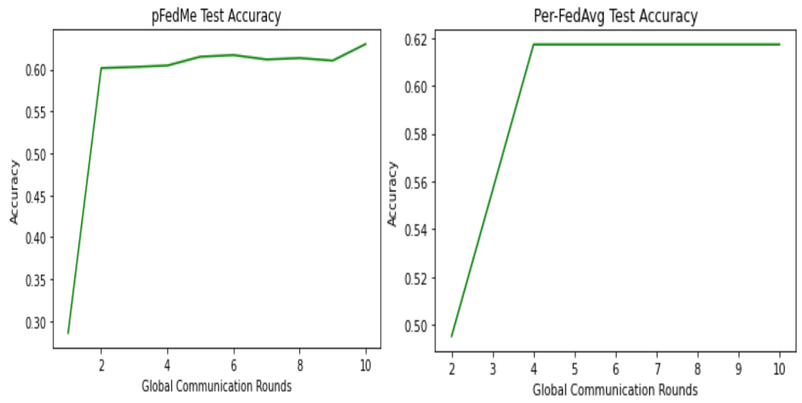
Results

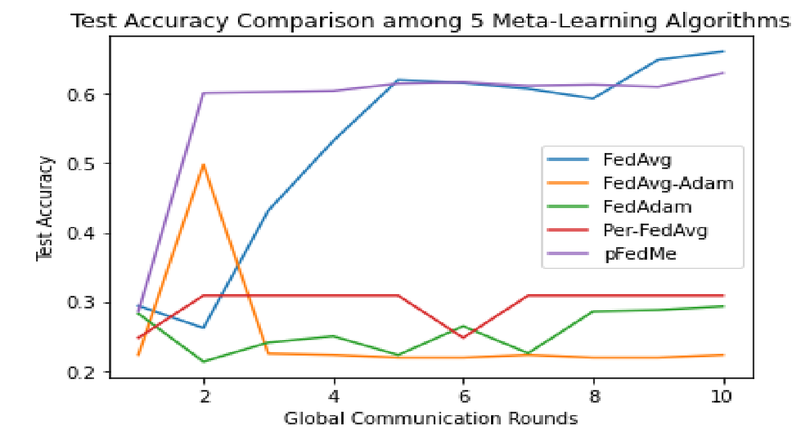
We trained and tested the 5 optimization algorithms under the same settings to ensure integrity. We split the data between 2 workers, in a non-iid format with a batch size of 16 and a client learning rate of 0.0001.











Conclusion

Efficiency is ensured whereas each and every user contributes to the training process with their full computational power. Every user obtains the global model and personalizes it based on their local data. Finally, privacy is maintained as the training happens in a federated setting. The raw data remains local and is never shared with the server.

References

1. G. Kaissis, A. Ziller, J. Passerat-Palmbach, T. Ryffel, D. Usynin, A. Trask, I. Lima, J. Mancuso, F. Jungmann, M.-M. Steinborn, et al., "End-to-end privacy preserving deep learning on multi-institutional medical imaging," Nature Machine Intelligence, vol. 3, no. 6, pp. 473–484, 2021.
2. D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," Cell, vol. 172, no. 5, pp. 1122–1131, 2018.
3. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778, 2016.
4. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial intelligence and statistics, pp. 1273–1282, PMLR, 2017.
5. S. Reddi, Z. Charles, M. Zaheer, Z. Garrett, K. Rush, J. Kone˘cn`y, S. Kumar, and H. B. McMahan, "Adaptive federated optimization," arXiv preprint arXiv:2003.00295, 2020.
6. A. Fallah, A. Mokhtari, and A. Ozdaglar, "Personalized federated learning with theoretical guarantees: A model-agnostic meta-learning approach," Advances in Neural Information Processing Systems, vol. 33, pp. 3557–3568, 2020.
7. C. T Dinh, N. Tran, and J. Nguyen, "Personalized federated learning with moreau envelopes," Advances in Neural Information Processing Systems, vol. 33, pp. 21394–21405, 2020.
8. J. Mills, J. Hu, and G. Min, "Multi-task federated learning for personalised deep neural networks in edge computing," IEEE Transactions on Parallel and Distributed Systems, vol. 33, no. 3, pp. 630–641, 2021.
9. J. Cheng, "brain tumor dataset," 4 2017.

Prepared for Thesis Poster Display Conference

11th -12th June 2022

Faculty of Management Technology
THESIS Poster Display Conference 2022