FedTMOS: Efficient One-Shot Federated Learning with Tsetlin Machine

Shannon How Shi Qi*, Jagmohan Chauhan†, Geoff V Merrett‡, Jonathan Hare‡
University of Southampton,
s.how@soton.ac.uk*, {gvm,jsh}@ecs.soton.ac.uk‡
University College London
jagmohan.chauhan@ucl.ac.uk†

I. Introduction

One-Shot Federated Learning (OFL) has emerged as a promising alternative to address the communication bottleneck. OFL restricts communication to a single round, thus minimizing communication errors, cost and reducing the risk of interference caused by iterative updates [1].

Current OFL methods that rely on Knowledge Distillation (KD) and ensemble learning aggregate local models into an ensemble before distilling it into a global model. A key challenge with these methods is their dependence on public datasets, which may be unsuitable for certain tasks [1]. Data-free methods with generative models [2], suffer from additional computational overhead. Neuron matching and model fusion techniques, eliminates the need for server-side training but struggle with performance when models are trained on heterogeneous data distributions due to misalignment of models [3]. On the client side, existing methods rely on Deep Neural Networks (DNNs), which are resource-intensive and impractical for clients with limited computational capabilities, such as edge devices. Therefore, we proposed a solution based on the Tsetlin Machine (TM) for efficient OFL [4], [5].

A. FedTMOS: One-Shot FedTM

Given J clients, each having local datasets $D_1, D_2, ..., D_J$. The objective is to aggregate the local TM models, $\mathbf{T} = \{T_1, T_2, ..., T_J\}$, into ϕ server models ($\phi < J$) that generalizes well over $\mathbf{D} \equiv \cup_{i \in \mathcal{J}} D_i$ in one communication round.

As illustrated in Figure 1, our method consists of two stages: first the weights are scaled based on the average normalized Gini index of all clients, followed by k-means clustering. Then, we perform a greedy reassignment of model weights to ϕ number of models to minimize overlap.

Our approach is intuitive: by maximizing inter-class separation within models, we enhance the model's ability to distinguish between classes. The Gini index quantifies data distribution to ensure balanced client participation by dynamically adjusting weights, ensuring fairness across clients [6].

II. Results

We compared our method with several FL algorithms representing different approaches for model aggregation. We demonstrate that FedTMOS outperforms all baseline methods in all settings by at least an average of 7.22% and the best data-free method by 12.79%. Furthermore, it achieves a reduction in upload communication costs by at least $2.3\times$, making FedTMOS well-suited for FL with edge devices and providing a strong foundation for further exploration into enhancing the efficiency and performance of FL.

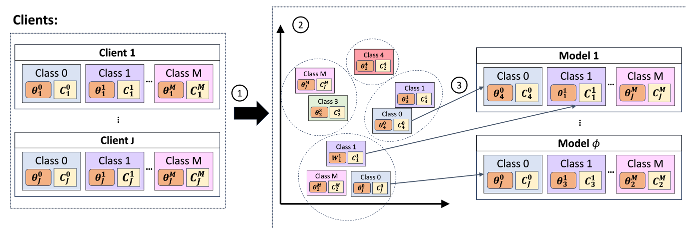

Fig. 1: Overview of FedTMOS: ① Clients upload their scaled clause weights, state parameters, and normalized Gini Index to the server. ② The server then rescales the weights using the mean normalized Gini Index and performs k-means clustering on the weights. This clustering is essential for grouping similar weights, helping to prevent large disparities in class weights across models during the reassignment process while also reducing complexity. ③ Finally, the number of models, denoted by a user-defined parameter ϕ, is initialized, and class weights from each cluster are greedily reassigned to maximize inter-class separation within each model.

III. Conclusion

FedTMOS is a novel framework for OFL that eliminates the need for server-side training and the use of synthetic or generated datasets. FedTMOS reduces computational complexity while achieving SOTA performance across heterogeneous settings.

References

[1] N. Guha, A. Talwalkar, and V. Smith, "One-Shot Federated Learning," 2019. [Online]. Available: https://arxiv.org/abs/1902.11175

[2] R. Dai, Y. Zhang, A. Li, T. Liu, X. Yang, and B. Han, "Enhancing One-Shot Federated Learning Through Data and Ensemble Co-Boosting," in *International Conference on Learning Representations*, 2024.

[3] D. Jhunjhunwala, S. Wang, and G. Joshi, "FedFisher: Leveraging Fisher Information for One-Shot Federated Learning."

[4] O.-C. Granmo, "The Tsetlin Machine - A Game Theoretic Bandit Driven Approach to Optimal Pattern Recognition with Propositional Logic," 2021.

[5] S. S. Q. How, J. Chauhan, G. V. Merrett, and J. Hare, "FedTM: Memory and Communication Efficient Federated Learning with Tsetlin Machine," in *2023 International Symposium on the Tsetlin Machine*, 2023, pp. 1–8.

[6] X. Li, S. Zhao, C. Chen, and Z. Zheng, "Heterogeneity-aware fair federated learning," *Information Sciences*, vol. 619, pp. 968–986, 2023.