FEDERATED LEARNING AGGREGATION VIA A REINFORCEMENT LEARNING POLICY

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent advances in federated learning have integrated an aggregation control policy trained with reinforcement learning. A research gap exists evaluating the performance impact of federated network elements as the reinforcement learning environment. This is particularly relevant for applications of machine learning in global health, which make use of federated learning to overcome cross-institution data-sharing constraints. In this work, we introduce a modular architecture of federated learning as a reinforcement learning environment. We conduct an experimental evaluation of policies trained in architecture configurations using a federated non-IID dataset and two deep reinforcement learning algorithms. Results of experiments show that choices of federated network elements only have a small effect on absolute classification accuracy (highest is 72.01%) for non-IID data, apart from the action aggregation strategy which is much lower. Findings are consistent with recent experiments, and this work provides a sandbox for robustly evaluating reinforcement learning methods in federated learning.

1 INTRODUCTION

Federated learning (FL) facilitates privacy-preserving, cross-institution data integration (McMahan et al., 2017). This enables regulatory-compliant machine learning at the scale of global health challenges. Applications in healthcare informatics include disease prediction, hospitalisation prediction, and adverse drug reaction prediction (Xu et al., 2021). Applied FL projects include the Federated Tumour Segmentation (FeTS) initiative (Baid et al., 2021) where thirty international healthcare institutions are jointly improving tumour boundary detection. Notable FL research includes detecting COVID-19 in chest CT scans using data from seven multinational hospitals (Dou et al., 2021).

A key challenge in FL is performance degradation in scenarios where data are not independent and identically distributed (non-IID) (Zhu et al., 2021). This challenge is common in health applications where data are often heterogeneous.

A trained reinforcement learning (RL) policy for FL aggregation has recently been demonstrated to improve FL performance amongst non-IID data (Nguyen et al., 2022). Authors configure the RL environment using the following FL elements: State = FL client loss, Reward = FL global model loss, Action = indirect, and Algorithm = DDPG (Lillicrap et al., 2015). Here we focus on understanding how the choice of FL elements as the RL environment affects performance. Identifying the comparative expressive power of FL elements can help researchers design optimal FL algorithms across global health contexts.

Our contributions are as follows:

1. We introduce a modular architecture for FL as an RL environment, with variable element designs influenced by federated optimisation literature.
2. We compare the performance of policies trained within modular architecture configurations, including against a baseline federated optimisation algorithm. In experiments, we discover limited performance impact under non-IID scenarios.
3. We release a public IPython Notebook as a sandbox for configurable training, deployment, and evaluation of our modular architecture.

2 RELATED WORK

Reinforcement Learning within Federated Learning. RL has been applied within federated networks to improve classification accuracy in a non-IID data setting (Nguyen et al., 2022), improve computation costs across a federated network (Zhan & Zhang, 2020), and improve federated network convergence speed (Wang et al., 2020). There is a remaining gap in knowledge related to the choice of architecture of FL as an RL environment.

Federated Optimisation Methods. Existing approaches to the non-IID data challenge include aggregating with consideration of model parameter divergence (Li et al., 2020), client variance (Li et al., 2019), batch normalisation (Li et al., 2021), or client personalisation (Fallah et al., 2020).

3 PRELIMINARIES

Federated Learning - A federated optimisation algorithm aims to find the minimum average sum of local loss functions across all training client models in the federated network: $\min_w \ f(w) = \frac{1}{n}\sum_{k=1}^{N} F_k(w)$, where the loss function of the k^{th} client is $F_k(w) = \mathbb{E}_{x_k \sim D_k}[f_k(w; x_k)]$, N is the number of clients, and D_k is the k^{th} client's data distribution.

Reinforcement Learning - An agent aims to learn an optimal control strategy for a dynamic environment. Through a repeated cycle of environment state observations S, taken actions A and received rewards R, following a finite Markov decision process, the agent's learning algorithm optimises a policy π to maximise the expected cumulative received reward.

4 METHODS

We present a modular architecture of FL as an RL environment (Figure 1). A RL agent will optimise its policy π by selecting an output across an action space of FL client aggregation factors $A_t^{(i..n)}$. The agent's goal is to maximise total episode reward R_{t+1}, which is received at each federated training round t as a mean environment indicator capturing the overall performance of network clients.

$$\text{Agent Goal: } \max_{\pi} \mathbb{E}_{\pi}\left[\sum_{k=t+1}^{T} \gamma^{k-t-1}\left(\frac{1}{n}\sum_{i=1}^{N} F_i(w)_t\right)_k\right]$$

where the RL reward $R_t = \frac{1}{N}\sum_{i}^{N} F_i(w)$ is discounted by factor γ^{k-t-1} to form a discounted sum of future rewards up to the terminal episode T.

We also introduce the architecture (Figure 2) for a configuration-consistent trained policy π to act as an aggregator within a FL network.

Within the modular architectures are the following configurable parameters:

State:

1. *Local client loss* captures the error between each federated network client model's classification labels and true labels, amongst local training data instances.
2. *Local client accuracy* captures the percentage of data instances, from a local validation dataset, which are correctly classified by the local model.
3. *Local client parameter divergence* captures the magnitude between the local model's parameters and the global federated model, motivated by FedProx (Li et al., 2020).
4. *Include local client dataset size ($|D_i|$)* is configurable parameter for inclusion alongside any of the above options, motivated by FedAvg (McMahan et al., 2017).

Action Aggregation Strategy:

1. *Direct aggregation strategy* where client parameters $w_t^{(i..n)}$ are vector-wise multiplied by the RL agent's selected actions $A_t^{(i..n)}$ to produce the federated network's shared global model w^T.

2. *Indirect aggregation strategy* where each client's parameters w_t^i are multiplied by the factor of the RL agent's produced influence actions $\frac{A_t^i}{\sum_i^n A_t^i}$, and w^T is the sum of these products.

Reward:

1. *Negative mean client loss* amongst all clients selected for training in the federated network.
2. *Mean client accuracy* amongst all clients selected for training in the federated network.

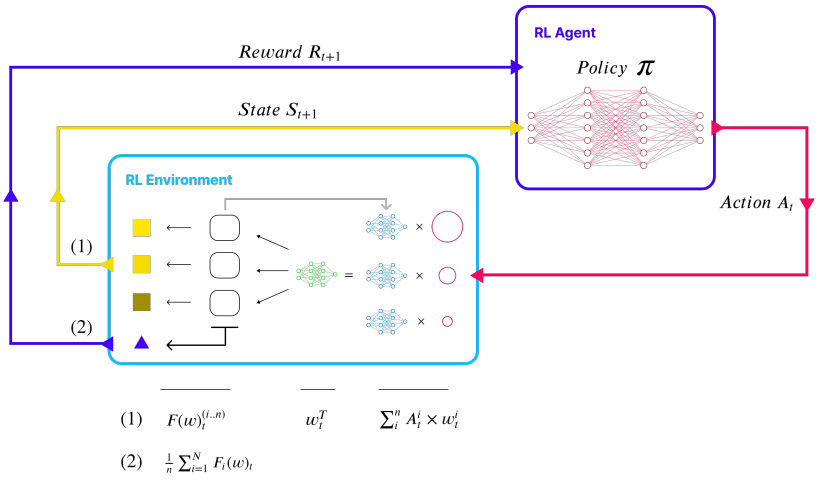

Figure 1: Schematic: FL as an RL environment.

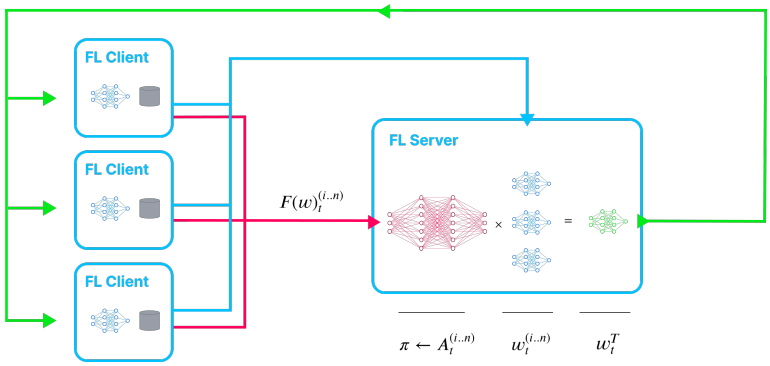

Figure 2: Schematic: Trained RL policy as a federated network aggregator.

5 Experiments & Results

We run experiments on Federated EMNIST, a FL dataset within the LEAF Benchmark for Federated Learning (Caldas et al., 2018). The image dataset has 3,550 unique clients, each with a data distribution across 62 label classes. Our implemented federated model is a convolutional neural network (CNN) for image classification that mirrors the implementation in 'Adaptive Federated Optimization' (Reddi et al., 2020). We evaluate the architectures on two deep RL algorithms, one On-Policy, PPO-Clip (Schulman et al., 2017), and one Off-Policy, Twin Delayed DDPG (TD3) (Fujimoto et al., 2018).

Our experiment design is a process of iterative variation to efficiently discover element expressive power (Table 1). The best-performing combination (72.01% accuracy) was made up of State = divergence, Reward = loss, $|D_i|$ = no, Action = indirect, and Algorithm = TD3. Configuration

variations had a limited impact on classification accuracy, other than action aggregation strategy (6.25% accuracy) (Figure 3). FedAvg benchmark performance across the 5 experimental stages had a mean classification accuracy of 71.51%.

Table 1: Experiment results: Absolute classification accuracy over configuration variations.

Stage	Experiment	State	Reward	$\|D_i\|$	Action	Algorithm	Accuracy
S1	E1	**Div**	Acc	No	Ind	TD3	71.67
	E2	Acc	Acc	No	Ind	TD3	71.05
	E3	Loss	Acc	No	Ind	TD3	71.44
S2	E4	**Div**	**Loss**	No	Ind	TD3	**72.01**
S3	E5	**Div**	**Loss**	**Yes**	Ind	TD3	70.68
S4	E6	**Div**	**Loss**	**No**	**Dir**	TD3	6.25
S5	E7	**Div**	**Loss**	**No**	**Ind**	**PPO**	71.05

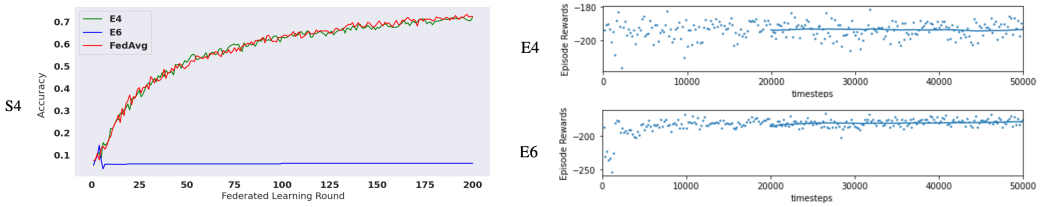

Figure 3: Stage Four evaluation results. E6 shows RL agent received reward across training. S4 shows E6, E4, and a FedAVG benchmark classification accuracy.

6 Discussion

Our results demonstrate that a trained RL policy is a robust aggregator in a federated network. The exception to this conclusion is the choice of direct aggregation strategy which fails to reasonably train a policy (Figure 3). The agent objective increases in complexity, from understanding how clients comparatively inform the reward to individually inform the reward. This result is consistent with the motivation for Multi-agent RL (MARL) in a federated network (Zhang et al., 2022).

While combinations of natural machine learning performance indicators yield similar evaluation results, the variation across element choice indicates further room for optimisation. A recommended research pathway for disease agnostic FL is to evaluate blended combinations of the performance elements explored in this study. To support this research, we have made an IPython Notebook which we make available for use by the community: *link redacted for anonymity.*

A limitation of this work is the evaluation across only one federated dataset. It is plausible that different FL elements as an RL environment will have distinctly varied expressive power across increasingly heterogeneous data. To evaluate this hypothesis, we recommend modular architecture experimentation on a synthetic dataset with engineered heterogeneity.

7 Conclusion

FL should be useful for global health problems where machine learning is used but data privacy needs to be preserved. RL can be used to support FL where data are non-IID, and experiments presented for one standard dataset show that a configuration using State = divergence, Reward = loss, $|D_i|$ = no, Action = indirect, and Algorithm = TD3 produces the highest performance. We present an implementation of a modular architecture for testing RL methods in FL and release the code and environment. Further experiments are needed to understand the impact of data heterogeneity across FL clients.

REFERENCES

Ujjwal Baid, Sarthak Pati, Siddhesh Thakur, Brandon Edwards, Micah Sheller, Jason Martin, and Spyridon Bakas. The federated tumor segmentation (fets) initiative. In *Neuro-Oncology*, volume 23, pp. 135–135. Oxford University Press Inc Journals Dept, 2021.

Sebastian Caldas, Sai Meher Karthik Duddu, Peter Wu, Tian Li, Jakub Konečnỳ, H Brendan McMahan, Virginia Smith, and Ameet Talwalkar. Leaf: A benchmark for federated settings. *arXiv preprint arXiv:1812.01097*, 2018.

Qi Dou, Tiffany Y So, Meirui Jiang, Quande Liu, Varut Vardhanabhuti, Georgios Kaissis, Zeju Li, Weixin Si, Heather HC Lee, Kevin Yu, et al. Federated deep learning for detecting covid-19 lung abnormalities in ct: a privacy-preserving multinational validation study. *NPJ digital medicine*, 4 (1):60, 2021.

Alireza Fallah, Aryan Mokhtari, and Asuman Ozdaglar. Personalized federated learning: A meta-learning approach. *arXiv preprint arXiv:2002.07948*, 2020.

Scott Fujimoto, Herke Hoof, and David Meger. Addressing function approximation error in actor-critic methods. In *International conference on machine learning*, pp. 1587–1596. PMLR, 2018.

Tian Li, Maziar Sanjabi, Ahmad Beirami, and Virginia Smith. Fair resource allocation in federated learning. *arXiv preprint arXiv:1905.10497*, 2019.

Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems*, 2:429–450, 2020.

Xiaoxiao Li, Meirui Jiang, Xiaofei Zhang, Michael Kamp, and Qi Dou. Fedbn: Federated learning on non-iid features via local batch normalization. *arXiv preprint arXiv:2102.07623*, 2021.

Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*, 2015.

Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pp. 1273–1282. PMLR, 2017.

Nang Hung Nguyen, Phi Le Nguyen, Duc Long Nguyen, Trung Thanh Nguyen, Thuy Dung Nguyen, Huy Hieu Pham, and Truong Thao Nguyen. Feddrl: Deep reinforcement learning-based adaptive aggregation for non-iid data in federated learning. *arXiv preprint arXiv:2208.02442*, 2022.

Sashank Reddi, Zachary Charles, Manzil Zaheer, Zachary Garrett, Keith Rush, Jakub Konečnỳ, Sanjiv Kumar, and H Brendan McMahan. Adaptive federated optimization. *arXiv preprint arXiv:2003.00295*, 2020.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Hao Wang, Zakhary Kaplan, Di Niu, and Baochun Li. Optimizing federated learning on non-iid data with reinforcement learning. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*, pp. 1698–1707. IEEE, 2020.

Jie Xu, Benjamin S Glicksberg, Chang Su, Peter Walker, Jiang Bian, and Fei Wang. Federated learning for healthcare informatics. *Journal of Healthcare Informatics Research*, 5(1):1–19, 2021.

Yufeng Zhan and Jiang Zhang. An incentive mechanism design for efficient edge learning by deep reinforcement learning approach. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pp. 2489–2498, 2020. doi: 10.1109/INFOCOM41043.2020.9155268.

Sai Qian Zhang, Jieyu Lin, and Qi Zhang. A multi-agent reinforcement learning approach for efficient client selection in federated learning. *arXiv preprint arXiv:2201.02932*, 2022.

Hangyu Zhu, Jinjin Xu, Shiqing Liu, and Yaochu Jin. Federated learning on non-iid data: A survey. *Neurocomputing*, 465:371–390, 2021.