Are All Users Treated Fairly in Federated Learning Systems?

Umberto Michieli\textsuperscript{1,2} and Mete Ozay\textsuperscript{1}

\textsuperscript{1} Samsung Research UK \textsuperscript{2} University of Padova

Abstract

Federated Learning (FL) systems target distributed model training on decentralized and private local training data belonging to users. Current methods aggregate models with importance proportional to the frequency of local samples. However, this leads to unfair aggregation with respect to users. Indeed, users with few local samples are considered less during aggregation and do not offer a real contribution to optimization of the models. In the real-world, statistical heterogeneity (e.g., highly imbalanced and non-i.i.d. data) is diffused and seriously harms model training.

We empirically analyze the relationship between fairness of aggregation of user models, accuracy of aggregated models and convergence rate of FL methods. We compare the standard FedAvg against a fair (uniform) scheme, i.e., FairAvg on benchmark datasets. Experimentally, we show that fair aggregation can be beneficial for accuracy and convergence rate, whilst reducing at the same time fluctuations of accuracy of the aggregate model when clients observe non-i.i.d. data.

Our Approach (FairAvg)

In general, federated aggregation can be expressed as:

\[
\sum_{k=1}^{K} \frac{w_k}{K} \theta_k
\]

FairAvg: \( a_k = 1/K \) (constant)

\( \rightarrow \) To show effect of data imbalance and non-i.i.d.-ness across clients

Federated Attention \( a_k \)

Note1: we subsample \( K = 10 \) clients

Note2: \( a_k \) computed by FedAvg is \( a_k \propto n_k \)

Many clients contribute little

Few clients dominate the scene

If data is highly non-i.i.d. this represents a problem for convergence

Results

ACF: AutoCorrelation Function

\( \rightarrow \) FairAvg improves accuracy wrt FedAvg (when data are distributed non-i.i.d.)

\( \rightarrow \) FairAvg reduces fluctuation towards convergence values wrt FedAvg (when data are distributed non-i.i.d.)

Conclusion

✓ We explore the relationship between fairness of aggregation schemes, accuracy of aggregated models and convergence rate of federated optimization methods on non-i.i.d. data.

✓ FairAvg is beneficial compared to FedAvg for both final accuracy and convergence rate, whilst reducing at the same time fluctuations of accuracy of the aggregate model.

✓ We believe that FL models could employ federated aggregation values centered around the value employed by FairAvg for uniform treatment of user contributions.

* Researched during internship at Samsung Research UK