Impact of non-IID data on the performance and fairness of differentially private federated learning

Saba Amiri, Adam Belloum, Eric Nalisnick, Sander Klaus, Leon Gommans

University of Amsterdam
Informatics Institute

Issues with Distributed ML in Medical Domain

Privacy

Differential Privacy

Consider adjacent datasets $A, B$ in $D$ which only differ in one element. The randomized mechanism $M: D \rightarrow R$ is $(\epsilon_1, \epsilon_2)$-differentially private if for any subset of outputs of $M$, $S \subseteq R$

$$Pr[M(A) \in S] \leq e^{\epsilon_2}Pr[M(B) \in S] + \delta.$$ (1)

where $\delta$ is the privacy budget, setting the level of intended privacy. The lower $\epsilon$, the higher the privacy level. $\delta$ is a small probability of failure of the DP guarantee. As a rule of thumb, it is set as less than $1/n$.

Differentially private SGD

1. Clip gradients
2. Add calibrated noise

Federated Learning

Algorithm 1: Federated Averaging

Data distribution: $\{A_1, A_2, \ldots, A_n\}$

1. Model initialization
2. Training samples $X_i$, labels $Y_i$, hyperparameters $\theta$. Round $t$,
3. Each client $i$ updates their model $\theta^{t}(x_i)$, training on $X_i, Y_i$. Model $\theta^{t}(x_i)$ does not share with the server,
4. Server aggregates $\theta^{t}_{avg}$.
5. If $t < T$ go to 2.

Data distribution: IID, non-IID, non-IID data

Differential Privacy

Utility Metrics

$q$ Precision
$q$ Recall
$q$ F1-Score

Fairness Metrics

$q$ Differential Fairness
$q$ Generalized Entropy Index
$q$ Equal Odds Rate

Data Distribution

SID assumption

$\Box$ Real world data distribution is non-IID

- Class imbalance: imbalance in target feature
- Feature imbalance: imbalance in non-target feature
- Node imbalance: imbalance in distribution of samples among nodes

Experimental Setup

Dataset

- Census Adult Income dataset
- Income as the target feature, "$\geq 50k$" as desirable outcome
- Race as the protected feature, "White" as privileged group

Fairness Metrics

$q$ Fairness drops with increase in privacy level
$q$ High privacy regimes act as a regularization method

Impact of non-IID data on dataset-level fairness

$q$ Fairness drops with increase in privacy level
$q$ Impact more prominent on more underprivileged groups
$q$ Non-IID distribution has a negative impact in low privacy regimes, impact less prominent with increase in privacy level

Impact of non-IID data on Group-Level Fairness

$q$ Fairness drops with increase in privacy level
$q$ Impact more prominent on more underprivileged groups
$q$ Non-IID distribution has a negative impact in low privacy regimes, impact less prominent with increase in privacy level

Impact of non-IID data on Performance

$q$ Performance drops with increase in privacy level
$q$ Recall drops significantly while the difference in precision is prominent but negligible
$q$ High privacy regimes act as a regularization method

More about EPI project: \url{https://enablingpersonalizedinterventions.nl}

Main references: