# Optimality and Stability in Federated Learning

Lokesh Mohanty

Department of Computational and Data Sciences
Indian Institute of Science

GTMD 2023

# Why is Federated Learning important?

- Local learning with insufficient data leads to high variance
- Transferring data from multiple locations is expensive
- Data privacy can also be a huge concern
- Solution: Share the parameters obtained by local learning and take their weighted average to get the global parameters

## Example (Group of Hospitals)

- Data: Patients information
- Problem: Diagnosing the patients
- Issue:
  - Patients data at a single hospital might not be sufficient due the problem having huge number of parameters
  - Patients information cannot be shared with others

[0] Li et al , Federated learning: Challenges, methods, and future directions, IEEE 2020

# Limitations of Federated Learning

- In federated learning, all the agents need not have the same underlying distribution
- This causes an increase in bias on joining a federation which leads to bias/variance trade-off
- Hasan[1] formulated this as a hedonic game where the agents can choose to join a federation
- Hedonic game allows to analyze the incentives of agents to contribute resources towards federated learning
- This study considers federated learning as a hedonic game

[1] Cengis Hasan, Incentive mechanism design for federated learning: Hedonic game approach, arXiv 2021

# Overview

- To analyze optimality and stability of federated learning, we need to make some assumptions on the agents and also we require a mathematical model of federated learning.
- Donahue[1] formulated a model for error in federated learning and the stability of federated learning
- Donahue[2] derived an algorithm to find an optimal arrangement and also analyzed optimality of federated learning
- Donahue[2] then related stability and optimality by giving bounds for Price of Anarchy(PoA)
- Then we will go over some studies related to our project.
- We will then conclude by stating the limitations and future directions.

[1] Kate Donahue and Jon Kleinberg, Model-sharing games: Analyzing federated learning under voluntary participation, AAAI 2021

[2] Kate Donahue and Jon Kleinberg, Optimality and Stability in Federated Learning: A Game-theoretic Approach, NeurIPS 2021

# Assumptions on the agents

- The assumptions made by Donahue[1] on the agents are
  - There are a fixed number of agents participating in federated learning
  - Each agents has access to local data and can contribute resources towards learning a global model
  - Agents are rational and self-interested (minimize their own error, while contributing as little as possible)
  - Agents can communicate with each other to form coalitions, this allows them to increase their overall performance
  - Error rates of each agent are known to all other agents. This allows for accurate calculation of optimal player arrangements
  - All agents have equal bargaining power. This ensures that no single agent can dominate the coalition formation

[1] Kate Donahue and Jon Kleinberg, Optimality and Stability in Federated Learning: A Game-theoretic Approach, NeurIPS 2021

# Model of federated learning

- Donahue[1] analyzed 3 variations of federated learning

## Uniform federation

Single global federation, all agents join this global federation

$$\hat{\theta}^f = \frac{1}{\sum_{i=1}^{M} n_i} \sum_{i=1}^{M} \hat{\theta}_i . n_i$$

## Coarse-grained federation

Single global federation, agents have the option to not join this federation

$$\hat{\theta}_j^w = w_j . \hat{\theta}_j + (1 - w_j) . \frac{1}{N} \sum_{i=1}^{M} \hat{\theta}_i . n_i$$

## Fine-grained federation

Multiple federations, agents can choose to join any federation. Most generalized variation

$$\hat{\theta}_j^v = \sum_{i=1}^{M} v_{ji} \hat{\theta}_i$$

[1] Kate Donahue and Jon Kleinberg, Model-sharing games: Analyzing federated learning under voluntary participation, AAAI 2021

## Stability

- Donahue[1] framed federated learning through the lens of cooperative game theory
- Each player wants to minimize their expected Mean-Squared-Error $err_i(C_i)$
- Derived exact terms for $err_i(C_i)($
- Gave conditions for which arrangements will be individually stable

[1] Kate Donahue and Jon Kleinberg, Model-sharing games: Analyzing federated learning under voluntary participation, AAAI 2021

# Optimality Model

C $\rightarrow$ Coalition, $\theta_i$ $\rightarrow$ Parameters of agent $i$, $n_i$ $\rightarrow$ Data size of agent $i$

## Error of a coalition for a particular agent

$$err_j(C) = \frac{\mu_e}{\sum_{i \in C} n_i} + \sigma^2 . \frac{\sum_{i \in C, i \neq j} n_i^2 + \left(\sum_{i \in C, i \neq j} n_i\right)^2}{\left(\sum_{i \in C} n_i\right)^2}$$

## Parameter estimation of a coalition

Weighted average of parameters of all agents in the coalition

$$\hat{\theta}_C = \frac{1}{\sum_{i \in C} n_i} . \sum_{i \in C} n_i . \hat{\theta}_i$$

## Optimality

Minimizes weighted average of errors across all agents

$$\frac{1}{\sum_{i=1}^{M} n_i} . \sum_{i=1}^{M} n_i . err_j(C_i)$$

# Optimality

- It is first shown that either local learning or federation can be arbitrarily far from the optimal. This shows that optimal arrangement is not necessarily a trivial case

**Lemma**

$\forall \rho > 1$, there exists a setting where local learning gives average error more than $\rho$ times higher than optimal.

**Lemma**

$\forall \rho > 1$, there exists a setting where federating in the grand coalition gives average error more than $\rho$ times higher than optimal.

# Optimality

- An efficient algorithm for calculating an optimal arrangement is then derived

**Theorem (Algorithm)**

*Given a set of agents* $\{i\}$ *with data size* $\{n_i\}$

- *Start with every agent doing local learning*
- *Begin grouping agents in ascending order of size*
- *Stop when the first player would increase its error by joining*
- *Resulting arrangement is optimal*

# Optimality

- The algorithm is proved by using building block lemmas which allows us to move from an arbitrary arrangement to an optimal one, in a cost-reducint way.
- This shows the correctness of the algorithm for calculating an optimal arrangement

**Building block lemmas**

- Stability
- Swapping
- Monotonicity of joining
- Monotonicity of leaving
- Merging

# Price of Anarchy

There are two cases where the Price of Anarchy is 1 (i.e., the optimal arrangement is stable)

### Definition (Price of Anarchy)

Price of Anarchy describes trade-off

$$\text{Price of Anarchy (PoA)} = \frac{\text{Error of worst stable arrangement}}{\text{Error of optimal arrangement}}$$

### Lemma

When all players are sufficiently large, local learning is both optimal and stable

### Lemma

When all players are sufficiently small, federation in grand coalition is both optimal and stable

# Price of Anarchy

- To find its upper bound, we need to upper bound the error of worst stable arrangement and lower bound the error of optimal arrangement
- From our assumptions we know that an agent's error is upper bounded by its local learning
- With the help of sub-lemmas relying on each agent's data size and the size of coalition it is federating with, we prove the following theorem

**Theorem (Constant Bound)**

$$PoA = \frac{\textit{maximum cost individually stable partition}}{\textit{lowest cost (optimal) partition}} \leq 9$$

# Related Work

- Donahue and Kleinberg[1] studied models of fairness
- Hu et al.[2] models clients behaviour in network
- Cui et al.[3] tries to find collaboration equilibrium
- Le et al.[4] analyzes incentives for agents to contribute computational resources while using an auction approach

[1] Kate Donahue and Jon Kleinberg, Models of fairness in federated learning, CoRR 2021
[2] Hu et al., Federated Learning as a Network Effects Game. CoRR 2021
[3] Cui et al., Collaboration equilibrium in federated learning. CoRR 2021
[4] T. H. Thi Le et al., An Incentive Mechanism for Federated Learning in Wireless Cellular Networks: An Auction Approach. IEEE 2021

# Conclusion

- We deeply studied the paper by Donahue[1] which presents a game-theoretic approach to federated learning that can improve accuracy rates while providing certain guarantees around social good properties such as total error
- This is a theoritical study and hence the results might be differ from acutal practice.
- The optimality bound has some assumptions on the cost definition and what is optimal, changing these can result in different results
- The construction of the framework for optimality and stability can be helpful for designing more complex models of federation with different definitions of optimality like fairness and also different notions of cost

*— Thank You —*

[1] Kate Donahue and Jon Kleinberg, Optimality and Stability in Federated Learning: A Game-theoretic Approach, NeurIPS 2021