Convergence of FedAvg

- The original paper [1] does not have theory.

- Paper [2] proved FedAvg converges for IID data.

- Paper [3] is the first to prove FedAvg (with SGD) converges for non-IID data.

- Paper [4] proved FedAvg (with GD) converges for non-IID data.

References

1. McMahan and others: Communication-efficient learning of deep networks from decentralized data. In *AISTATS*, 2017.
2. Stich: Local SGD converges fast and communicates little. In *ICLR*, 2018.
3. Li and others: On the convergence of FedAvg on non-IID data. *arXiv*, 2019.
4. Khaled and others: First analysis of local GD on heterogeneous data. *arXiv*, 2019.