Measuring the Effects of Non-Identical Data Distribution for Federated Visual Classification

Tzu-Ming Harry Hsu, Hang Qi, Matthew Brown

stmharry@google.com, hangqi@google.com, mtbr@google.com

Motivation & Contributions

- In real-world Federated Learning, each participating user has very different and non-identical distribution.

- A realistic and practical setting for Federated Learning is needed to study the effect of non-identical data and the pool size of clients.

Synthetic Non-Identical Data

- We synthesize 100 federated learning clients from CIFAR-10 by:
  1. Draw class-marginal distribution from Dirichlet distribution $q \sim \text{Dir}(\alpha p)$, where $p$ is uniform.
  2. Assign $(500 \times q)_c$ examples from class $c$ for all classes.

Methods

- Federated Averaging (FedAvg) updates the weights via:
  1. Select a fraction $C$ of all clients to report.
  2. Locally train clients with their respective data for $E$ local epochs and yield local model updates $\{\Delta w_k\}_{k=1}^K$.
  3. Update server weights by $w \leftarrow w - \Delta w$, where $\Delta w = \sum_{k=1}^K \frac{\# \text{examples for client } k}{\# \text{total examples}} \Delta w_k$.

- Federated Averaging with Momentum (FedAvgM) updates server weights by $w \leftarrow w - \nu v + \Delta w$, where $v \leftarrow \beta v + \Delta w$.

Results

Non-Identical Data on FedAvg

Learning Curves

FedAvgM

Performance Curves

Hyperparam Sensitivity

References

Sattler et al. Federated and communication-efficient learning from non-IID data. arXiv, 2019.