

パーソナライズド連合学習手法の比較と分析

松田 光司[†] 佐々木 勇和[‡] 肖 川[§] 鬼塚 真[¶]
 大阪大学[†] 大阪大学[‡] 大阪大学[§] 大阪大学[¶]

1 はじめに

連合学習はサーバと複数のクライアントがデータを共有することなく、協調して機械学習モデルを構築する分散型の機械学習手法である。連合学習における課題の一つとして、データの不均一性の問題がある。各クライアントが持つデータの分布は異なるため、各クライアントに対して最適な単一のグローバルモデルを学習することは困難である。データの不均一性に対処するために、クライアント毎に最適なモデルを作成する、パーソナライズド連合学習手法が多く考案されている。しかし、これらの手法の網羅的な比較及び分析は行われていない。そこで、本稿では4つのデータセットを用いて、パーソナライズド連合学習手法の比較と分析を行い、手法の特性を明らかにする。

2 連合学習手法

本稿の実験で比較する手法について説明する。非パーソナライズド連合手法として FedAvg [1] と Fedprox [2]、パーソナライズド連合手法として HypCluster [3]、MAPPER [3]、FML [4]、FedMe [5]、LG-FedAvg [6]、FedPer [7]、FedRep [8]、Ditto [9] および pFedMe [10] を用いる。

非パーソナライズド連合学習手法: FedAvg [1] は連合学習の代表的な手法であり、サーバが一つのモデルを作成する。サーバは各クライアントが学習したモデルを受け取り、平均化して1つのモデルに集約する。Fedprox [2] は FedAvg を拡張した手法であり、各クライアントの学習の損失関数に近接項を追加することで、モデルパラメータの発散を防ぐ。

パーソナライズド連合学習手法: HypCluster [3] はサーバが複数のモデルを作成する。クライアントは自身のデータを用いて最も損失が小さいモデルを選択し、そのモデルを学習する。MAPPER [3] はサーバと各クライアントがそれぞれモデルを作成する。各クライアントは自身のモデルとサーバのモデルを加重平均し補間モデルを作成し、補間モデルを最適化するように自身のモデル、サーバのモデル、および加重平均の比率を学習する。

FML [4] と FedMe [5] はサーバと各クライアントがそれぞれモデルを作成する。FML は各クライアントは自

身のモデルとサーバのモデルを深層相互学習によって学習する。推論時はサーバのモデルを用いることもできるが、本稿では各クライアントが作成したモデルを用いる。FedMe は各クライアントが自身のモデルと他クライアントのモデルを交換し、互いのモデルを深層相互学習によって学習する。

LG-FedAvg [6]、FedPer [7]、および FedRep [8] はサーバがモデルの一部分を作成し、各クライアントがモデルの残りの部分を作成する。LG-FedAvg はサーバがモデルの出力側、クライアントがモデルの入力側の一部を作成するのに対し、FedPer と FedRep はサーバがモデルの入力側、クライアントがモデルの出力側の一部を作成する。また、LG-FedAvg と FedPer がモデルの出力側と入力側を同時に学習するのに対し、FedRep はモデルの出力側と入力側の片方のモデルパラメータを固定し、それらを交互に学習する。推論時には、各クライアントは自身のモデルの一部とサーバのモデルの一部を結合したものをを用いる。

Ditto [9] と pFedMe [10] はサーバと各クライアントがそれぞれモデルを作成する。各クライアントとサーバのモデルパラメータが離れすぎないように学習する。Ditto はサーバとクライアントのモデルパラメータの差分を各クライアントのモデルに加算するのに対し、pFedMe ではモロー包絡を用いてサーバのモデルと離れすぎないように正則化項を損失関数自体に組み込んでいる。

3 評価実験

実験では各手法の精度検証を行う。Pytorch を用いて単一の GPU マシンで仮想的にクライアントとサーバを作成する。

実験設定: データセットとして、Federated EMNIST-62 (FEMNIST)、MNIST、CIFAR-10、Shakespeare の4つを用いる。FEMNIST と MNIST では CNN を用いて、CIFAR-10 では VGG を用いて多クラス分類を行う。Shakespeare では LSTM を用いて次文字予測を行う。全てのデータセットでクライアント数は20とし、クライアントが保持するデータを変えて5回実験を行なった。

比較手法として、2章で説明した手法に加えて、各クライアントが自身のデータのみを使用して学習する Local Data Only も比較する。また、FedAvg、FedProx、HypCluster、FedMe はモデル構築後、それぞれのデータで fine-tuning する。

実験結果: 実験結果を表1にテストデータに対する精度の平均と標準偏差を表している。

A performance comparison and analysis of Personalized Federated Learning methods

[†] Koji Matsuda, Osaka University

[‡] Yuya Sasaki, Osaka University

[§] Chuan Xiao, Osaka University

[¶] Makoto Onizuka, Osaka University

表1 実験結果

	FEMNIST	MNIST	CIFAR-10	Shakespeare
Local Data Only	64.71 ± 2.94	97.00 ± 0.99	73.17 ± 1.55	24.77 ± 1.95
FedAvg	77.25 ± 3.99	99.07 ± 0.04	89.59 ± 0.94	42.53 ± 2.19
FedProx	76.96 ± 3.42	98.79 ± 0.06	89.76 ± 0.62	45.17 ± 2.83
HypCluster	76.29 ± 3.15	98.95 ± 0.11	88.54 ± 1.42	41.10 ± 3.29
MAPPER	60.95 ± 3.04	97.00 ± 0.47	61.29 ± 4.19	36.77 ± 1.58
FML	67.91 ± 2.53	98.16 ± 0.17	79.89 ± 1.44	28.73 ± 1.78
FedMe	78.06 ± 3.00	99.03 ± 0.14	90.96 ± 0.84	45.83 ± 2.48
pFedMe	72.92 ± 3.54	98.96 ± 0.05	79.46 ± 2.08	40.33 ± 2.27
LG-FedAvg	65.14 ± 3.12	97.80 ± 0.16	78.53 ± 1.57	23.17 ± 1.93
FedPer	65.96 ± 2.81	99.06 ± 0.04	90.00 ± 0.83	30.83 ± 3.32
FedRep	66.04 ± 2.20	98.90 ± 0.09	88.96 ± 0.48	31.71 ± 2.29
Ditto	75.68 ± 3.63	99.13 ± 0.05	90.41 ± 0.67	49.33 ± 1.85

表1より、まず多くの連合学習手法はlocalより精度が高い。各クライアントの持つデータのみでの学習では過学習を起こすため、連合学習手法の有効性がわかる。

全手法の中ではFEMNISTとCIFAR-10に対してはFedMeが、MNISTとShakespeareに対してはDittoがそれぞれ最も精度が高い。特に、FedMeでは他の2つのデータセットに対してもDittoと同等程度、あるいは2番目に精度が高いことからデータセットに対して堅牢であることがわかる。FedPerとFedRepは、MNISTとCIFAR-10に対しては最も精度が高い手法と同等程度の精度であり、FEMNISTとShakespeareに対してはLocal Data Onlyと同等程度の精度である。これは、FedPerとFedRepではモデルの出力側は各クライアント上のデータでのみ学習するため、各クライアントが保持するデータ数が小さいFEMNISTとShakespeareでは過学習を起こしてしまうためである。FedAvgとFedproxでは学習後にfine-tuningしており、全てのデータセットに対して最も精度が高い手法と同等程度の精度である。このことから、単純な連合学習手法とfine-tuningを組み合わせることで十分にデータの不均一性に対処できることがわかる。

謝辞 本研究はJSPS科学研究費JP17H06099およびJP20H00584の支援によって行われた。実験には産総研のAI橋渡しクラウド(ABCI)を利用した。

参考文献

- [1] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *AISTATS*, pp. 1273–1282, 2017.
- [2] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. In *MLSys*, Vol. 2, pp. 429–450, 2020.
- [3] Yishay Mansour, Mehryar Mohri, Jae Ro, and Ananda Theertha Suresh. Three approaches for personalization with applications to federated learning. *arXiv*, 2020.
- [4] Tao Shen, Jie Zhang, Xinkang Jia, Fengda Zhang, Gang Huang, Pan Zhou, Kun Kuang, Fei Wu, and Chao Wu. Federated mutual learning. *arXiv*, 2020.
- [5] Koji Matsuda, Yuya Sasaki, Chuan Xiao, and Makoto Onizuka. Fedme: Federated learning via model exchange. *arXiv preprint arXiv:2110.07868*, 2021.
- [6] Paul Pu Liang, Terrance Liu, Liu Ziyin, Ruslan Salakhutdinov, and Louis-Philippe Morency. Think locally, act globally: Federated learning with local and global representations. *arXiv preprint arXiv:2001.01523*, 2020.
- [7] Manoj Ghuhana Arivazhagan, Vinay Aggarwal, Aaditya Kumar Singh, and Sunav Choudhary. Federated learning with personalization layers. *arXiv preprint arXiv:1912.00818*, 2019.
- [8] Liam Collins, Hamed Hassani, Aryan Mokhtari, and Sanjay Shakkottai. Exploiting shared representations for personalized federated learning. In *ICML*, pp. 2089–2099, 2021.
- [9] Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through personalization. In *ICML*, pp. 6357–6368, 2021.
- [10] Canh T. Dinh, Nguyen Tran, and Josh Nguyen. Personalized federated learning with moreau envelopes. In *NIPS*, pp. 21394–21405, 2020.