データの不均一性を考慮した連合学習技術の比較と分析

松田 光司 † 堀 敬三 ‡ 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 ¶ 大阪大学 ‖

1 はじめに

近年では，スマートフォンやタブレットなどのデバイスの増加に伴い，これまでにないほどの大量の個人データの入手が可能である．それらのデータを用いた機械学習が様々なアプリケーションに応用されている．従来の機械学習では中央のサーバが全てのデータを保持しながら集中的にモデルを学習するため，サーバにデータを全て送る必要がある．しかし，デバイスから入手したデータは機密性やプライバシが高いという問題がある．加えて，ネットワーク帯域幅の制限があるため，全てのデータを送ることは現実的ではない．

これらの問題を解決するために連合学習が提案された [1]．連合学習は複数のクライアントが中央のサーバと連携して，クライアントの持つデータを共有することなくモデルを学習する分散型の機械学習技術である．しかし，クライアントのもつデータには不均一性がある．例えば，『I live in...』の後に続く単語はユーザによって異なる．このようなデータの不均一性がある場合，既存の連合学習手法ではモデルの推論性能が低下する．データの不均一性に対処するために様々な手法が提案されているが，網羅的な比較および分析は行われない．そこで，本稿では２つのデータセットを用いて，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 連合学習技術

本稿の実験で比較する 5 つ手法について説明する．FedAvg [1] は連合学習の代表的な手法である．まず，各クライアントがサーバからモデルをダウンロードし，自身のデータでモデルを学習する．各クライアントは学習後のモデルをサーバに送る．そして，サーバがそれらのモデルを平均化して一つのモデルに集約する．この手順をモデルが収束するまで繰り返す．

HypCluster [2] はモデルを用いたユーザクラスタリングを行う．HypCluster では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 [2] はサーバが一つのモデルを作成し，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する．そして補間モデルを最適化するように自身のモデル，ダウンロードしたモデル，加重平均の比率を学習する．

FedMD [3] は，各クライアントがそれぞれモデルを作成する．FedMD は（1）クライアント毎にパブリックデータで学習した後自身のデータで学習する転移学習と，(2）クライアントのモデルのパブリックデータに対する出力をサーバに送信し，平均化されたものを正解ラベルとする知識蒸留の 2 つのステップで学習を進める．パブリックデータとは，全てのクライアントがアクセスできるデータを指す．

FML [4] はサーバが一つのモデルを作成し，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深層相互学習によって学習する．本稿では，推論に各クライアントがそれぞれ作成したモデルを用いる．

3 評価実験

実験では，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し，精度検証を行う．

実験設定: データセットとして Federated EMNIST-62 データセット（FEMNIST）と Shakespeare データセットの 2 つを利用する．FEMNIST は 28×28 ピクセルの画像データであり，CNN を用いて多クラス分類を行う．Shakespeare データセットは，演劇中の役がぞれぞれクライアントに割り当てられ，LSTM を用いて演劇中のセリフの次に来る文字（計 80 種類）を推論する．

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

A performance comparison and analysis of Federated Learning methods for data heterogeneity

† Koji Matsuda, Osaka University
‡ Keizo Hori, Osaka University
§ Yuya Sasaki, Osaka University
¶ Chuan XIAO, Osaka University
‖ Makoto Onizuka, Osaka University

表 1　実験結果

手法	FEMNIST	Shakespeare
local	66.59	24.18
FedAvg	86.30	48.80
FedAvg(fine-tune)	**86.84**	48.90
HypCluster	86.31	**51.32**
HypCluster(fine-tune)	86.25	51.06
MAPPER	82.07	42.96
FedMD	71.09	41.10
FML	71.34	30.22
Centra	86.60	55.13
Centra(fine-tune)	87.23	55.29

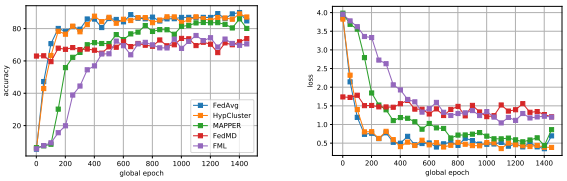

図 1　FEMNIST 学習曲線

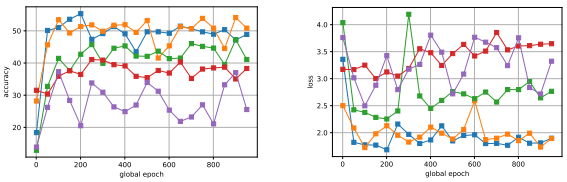

図 2　Shakespeare 学習曲線

習する local と全データをサーバが持っているものとしてモデルを学習する Centra を比較する．また，Centra，FedAvg，HypCluster は，各クライアントのデータで再学習する fine-tune も比較する．

実験結果: 実験結果を表 1，図 1，および図 2 に示す．表 1 は，最終的なモデルを用いたテストデータに対する各クライアントの精度の平均を表している．図 1 と図 2 は，エポック毎の参加クライアントの検証用データに対する精度の平均である．

表 1 より，まず連合学習手法は local より精度が高い．ローカルデータのみの学習では，各クライアントの持つデータのみ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．次に，連合学習手法の精度は Centra よりも低い．2 つのデータセットには不均一性があるためであり，連合学習における精度向上のためにはデータの不均一性に対応することが重要であることがわかる．

既存手法の中では FEMNIST と Shakespeare に対しては FedAvg（fine-tune）と HypCluster がそれぞれ最も精度が高い．HypCluster では，FEMNIST に対してはほとんど FedAvg と精度が変わらないが，Shakespeare に対しては FedAvg よりも精度が高い．これは FEMNIST よりも Shakespeare の方がデータの不均一性があり，ユーザクラスタリングの効果が高いためだと考えられる．FedMD と FML は，local よりも精度が高いが他の手法に比べて精度が低い．FedMD と FML の精度が低いのは，どちらもクライアントのモデルが各クライアントの持つデータ上でしか学習しないためである．

図 1 と図 2 より，まず Sakespeare の学習曲線が FEMNIST と比べて安定していない．これは Shakespeare の方がクライアント毎の精度に差があるためである．次に 2 つのデータセットで FedMD の初期精度が最も高い．これはパブリックデータを用いた転移学習をおこなっているためである．また，MAPPER の精度の収束が FedAvg と HypCluster に比べて遅い．これは MAPPER ではクライアントのモデル，サーバのモデル，加重平均の比率の 3 つを学習しているためである．図 1 と図 2 の結果より，収束速度，精度ともに FedAvg と HypCluster が他手法よりも優れていることがわかる．

謝辞 本研究は JSPS 科研費 JP20H00584 の助成を受けたものです。

参考文献

[1] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics*, pp. 1273–1282, 2017.

[2] Yishay Mansour, Mehryar Mohri, Jae Ro, and Ananda Theertha Suresh. Three approaches for personalization with applications to federated learning. *arXiv preprint arXiv:2002.10619*, 2020.

[3] Daliang Li and Junpu Wang. Fedmd: Heterogenous federated learning via model distillation. *arXiv preprint arXiv:1910.03581*, 2019.

[4] Tao Shen, Jie Zhang, Xinkang Jia, Fengda Zhang, Gang Huang, Pan Zhou, Kun Kuang, Fei Wu, and Chao Wu. Federated mutual learning. *arXiv preprint arXiv:2006.16765*, 2020.