연합 학습과 클라이언트 분포 추적

강동석
광주과학기술원 인공지능 대학원
dongseok176@gm.gist.ac.kr

안창욱
광주과학기술원 인공지능 대학원
cwan@gist.ac.kr

Federated Learning and Tracking Client Distribution

Dongseok Kang
Gwangju Institute of Science and Technology
AI Graduate School

Chang-Wook Ahn
Gwangju Institute of Science and Technology
AI Graduate School

Abstract

연합 학습은 여러 클라이언트에서 수집한 훈련 결과를 취합하여 하나의 공통된 모델에서 전역해를 찾으려 하는 방법이다. 학습에 참여하는 클라이언트들이 가지는 데이터는 알려지지 않은 분포에서 추출한다고 가정하는데, 서로 다른 데이터 분포로 인한 편향으로 글로벌 모델의 학습에 어려움이 있다. 이 문제를 해결하기 위해 클라이언트 분포를 추적하고 이를 글로벌 모델의 학습에 반영하는 방법을 고안하고 검증하였다. 클라이언트의 학습 결과를 취합할 때 클라이언트가 속한 그룹에 따라 업데이트되는 가중치의 크기를 조정하였다. 이미지 분류 실험에서 이 방법을 적용하였을 때 그렇지 않은 경우 대비 정확도의 향상을 확인할 수 있었다.

1. 서 론

데이터의 이동을 허용하지 않고 글로벌 모델을 학습시키는 연합 학습은 데이터 프라이버시를 지키면서 인공지능 모델을 학습할 수 있는 방법으로 각광받았다[1, 2]. 일반적인 경우 데이터를 한 곳에 취합하여 모델을 학습하는 반면에, 연합 학습은 데이터 대신 글로벌 모델의 가중치를 취합하여 학습한다. 이 때 클라이언트가 소유한 데이터를 공개하지 않고 지역적으로 학습하기 때문에 데이터 프라이버시를 지킬 수 있다.

그러나 관련 연구들은 개별 데이터들이 어디에도 공개되지 않기 때문에 글로벌 모델을 학습시키는 데에 어려움이 있음을 확인하였다. 개별 클라이언트가 가지는 데이터의 분포, 학습에 참여하는 기계의 시스템 특성 등이 제각각 달라서 발생하는 문제들은 글로벌 모델의 학습을 어렵게 했다. 특히 클라이언트들이 지역적으로 가지고 있는 데이터의 분포가 서로 매우 다르고, 일부 데이터들이 학습에 참여하는 경우가 낮을 때 글로벌 모델의 전역 최적해를 찾기가 쉽지 않다[3, 4]. 데이터 자체가 가지는 복잡도에 더해 분산된 클라이언트들로 인하여 발생하는 편향이 최적해를 찾기 어렵게 한다.

연학 학습에서 클라이언트의 데이터 분포는 알려지지 않기에, 모든 데이터와 클라이언트를 같은 분포에서 추출한다고 가정하고 학습 결과들을 산술평균으로 취합하는 것이 일반적이다[5]. 실제로는 데이터의 편향이 있을 것이고, 클라이언트 또한 서로 같지 않는 가용성을 가지기에 이는 전역해를 찾는 최적의 방법이 아니다.

우리는 클라이언트들을 그룹으로 나누어 그 분포를 추적하고, 학습 결과를 취합할 때 이를 사용하는 방법을 제안하고 검증하였다. 연합 학습에 참여하는 클라이언트의 그룹 분포를 통하여 각 그룹이 라운드마다 참여하는 비율을 계산하였다. 이 값을 통하여 모든 그룹이 동일한 비율로 글로벌 모델을 학습할 수 있도록 학습을 조율하였다. 이를 통하여 과적합된 결과의 크기를 줄이고, 비교적 덜 최적화된 데이터에 대한 결과의 크기를 키울 수 있었다. 이를 이미지 데이터에 대한 분류 학습 실험으로 확인할 수 있었는데, 우리가 사용한 방법을 적용했을 때 기 방법 대비 정확도의 향상을 확인할 수 있었다.

2. 관련 연구

연합 학습은 지역적으로 분산된 데이터를 학습하기 위해 글로벌 모델을 배포하고, 이를 취합하여 평균을 통해 학습시키는 FedAvg 방법이 대표적이다[5]. 평균을 통한 학습 방법은 다양한 데이터 환경에서 실험되었으나, Non-IID 데이터의 경우에는 그 성능의 열화가 확인되었다.

평균 방법의 변종으로 정규화를 통하여 학습 데이터를 취합하는 방법도 연구되었다[6]. 배포되는 글로벌 모델이 변화하는 정도를 정규화를 통하여 학습에 반영하고, 글로벌 모델과 지역 모델의 차이를 조절하여 데이터 학습의 일반화를 달성할 수 있었다.

평균을 통한 학습 결과의 취합과 정규화를 통한 학습 결과의 취합은 통신 비용과 일반화 능력을 상승시킬 수 있었으나, 데이터의 분포를 추적하지는 않았다. [7]은 학습에 참여하는 클라이언트를 서버가 선택하여 학습하는 방법을 연구하였다. 이 방법들은 클라이언트를 원하는 대로 선택하는 게 가능하다고 가정하였다. [8]는 딥 Q-러닝 알고리즘을 통하여 클라이언트 서브셋을 선택하고, 방법 5는 멀티 암드 밴딧을 통하여 서브셋을 선택한다. 서브셋을 선택할 때 클래스 분포는 모델의 그래디언트 업데이트를 프록시 데이터와의 차를 이용한 연구이다.

이 성과는 2022년도 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1A2C3013687).

3. 방법

연합 학습은 매 라운드마다 K개의 클라이언트를 추출하고, 각 클라이언트에서 공통 모델은 지역 데이터에 맞추어 $\boldsymbol{E}$번 학습한다. 클라이언트의 결과를 취합할 때는, 결과값의 산술 평균으로 취합하는 대신에 제안하는 방법을 사용한다.

우리는 각 클라이언트가 $\boldsymbol{G}$개의 그룹 중 하나에 속한다고 가정한다. 각 그룹은 서로 다른 데이터 분포를 가지고, 같은 그룹 내 클라이언트는 같은 데이터 분포를 공유한다. 연합 학습의 특성상 개별 데이터의 서로 다른 분포를 알 수는 없지만, 클라이언트를 추적하여 각 클라이언트가 추출되는 확률을 계산할 수 있다. 이를 사용한 알고리즘은 다음과 같다.

ALGORITHM: Federated Learning Algorithm Tacking Groups

$\theta_0 \leftarrow Initialize\ model\ parametr$

$p_0^k \leftarrow 1/K$

For $t = 0, \dots, T-1$ **do**

Server selects $\boldsymbol{K}$ devices from $\boldsymbol{G}$ groups

Server sends θ_t to all devices

Each device k updates θ_t for $\boldsymbol{E}$ epochs of **SGD**

Server aggregates θ_t^k from all devices

$$\theta_{t+1} \leftarrow \sum_{k \in K} (1-\alpha)\frac{\theta_t^k}{|K|} + \alpha p_t^k \theta_t^k$$

$$p_{t+1}^k \leftarrow \beta \cdot p_t^k + \frac{|\{client \mid client \in K\ ,and\ client\ is\ same\ group\ of\ k\}|}{|K|}$$

End For

매 라운드마다 각 그룹별 출현 빈도 p_t^k를 계산한다. 이 때 β를 모멘텀으로 사용하였다. 학습이 완료된 지역 모델 θ_t^k은 p_t^k 와의 곱으로 취합된다. 해당 취합 방법의 영향을 조절하기 위하여 α를 사용한다.

4. 실험

제안하는 방법의 효용성을 검증하기 위하여 CNN을 사용한 이미지 분류 모델을 연합 학습을 통하여 학습하였다. 모든 실험에서 동일한 CNN[9] 모델을 사용하였고, 이미지 데이터셋은 FashionMNIST[10]이다. 모든 실험은 이미지 분류 모델을 1e-2의 학습률로 100라운드, 1e-3의 학습률로 다시 100라운드를 학습하였다. 매 라운드마다 서버는 100개의 클라이언트를 그룹별 가중치를 가지고 추출하고, 클라이언트에서는 32배치 사이즈로 5라운드동안 학습하게 된다.

데이터가 클라이언트에게 분산된 환경을 시뮬레이션 하기 위하여 데이터셋의 훈련 데이터를 임의로 나누어 클라이언트에게 분배하였다. 데이터셋의 레이블을 3개의 그룹으로 나누고, 각 그룹마다 100개의 클라이언트를 생성하였다. 각 클라이언트는 500개의 이미지를 가지는데, 개별 이미지는 0.9의 확률로 클라이언트의 그룹에 속한 레이블을 갖는다.

위와 같이 생성한 그룹에서 서버는 100개의 클라이언트를 추출하는데, 각 그룹별로 0.6, 0.2, 0.2의 가중치를 가지고 클라이언트들을 추출하였다. β는 0.8을 사용하였다. 글로벌 모델의 훈련에는 60,000개의 이미지, 글로벌 모델의 검증에는 훈련에 사용되지 않은 테스트 이미지를 가지고 정확도를 관찰하였다.

총 4개의 실험을 진행하였는데, 각 실험에서 α를 $0, 0.05, 0.1, 0.2$로 맞추어 제안한 알고리즘의 영향을 관찰하였다.

5. 결과

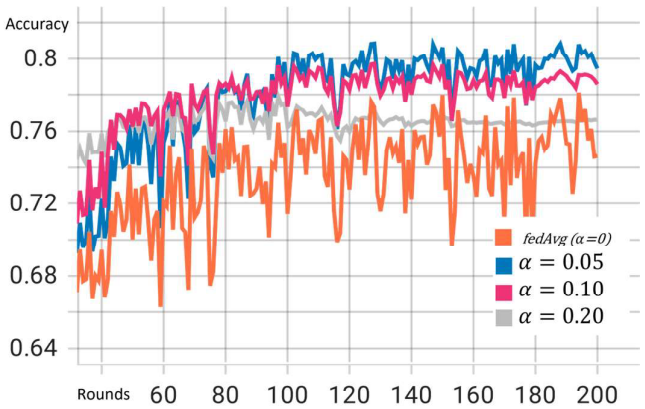

그림 1. α 의 변화의 따른 정확도

실험에서 제안하는 방법을 적용한 모든 경우에서, 그렇지 않은 경우에 대비해 정확도 상승을 관찰할 수 있었다. α가 0 일 때에 74%의 정확도, α가 0.05일 때 79% 정확도로 최대 5%의 향상을 확인하여 알고리즘의 효용성을 검증할 수 있었다.

정확도 상승과 더불어 제안하는 알고리즘은 학습이 더 안정적임을 관찰할 수 있었다. 이는 데이터가 편향되어 추출되었을 때, 모델의 가중치 변화가 특정 방향으로만 크게 바뀌는 것과 관련이 있다. 자주 추출되는 그룹의 클라이언트에서 얻은 가중치 크기를 줄임으로서 모델이 특정 방향으로 치우지지 않고 안정적인 방향으로 학습이 가능했다. 이는 α의 변화로 확인할 수 있는데, 더욱 큰 α를 사용하여 적극적으로 가중치 취합을 조절한 경우에 학습이 더 안정적임을 확인할 수 있었다.

모든 α에서 정확도의 향상을 관찰할 수 있었지만, α의 크기에 따라 정확도 향상 폭이 달랐다. α가 커질수록 오히려 정확도 향상폭은 줄어들었는데, 이는 가중치 크기의 조절만으로는 전역해를 찾는데 어려움이 있음을 보여준다.

모델의 가중치 공간에서 현재 그룹들이 가지는 가중치의 선형 조합은 각 데이터의 분포를 고려하지 않고 각 그룹이 추출되는 확률만을 고려하고 있기에 α 가 커질수록 전역해와의 괴리가 커졌다고 해석된다.

6. 결론

연합 학습에서 클라이언트의 데이터 특성에 따라 그룹으로 나뉜다고 가정하고, 이를 고려하여 가중치 취합에 활용할 경우 정확도 향상을 얻을 수 있음을 확인할 수 있었다.

이 성과는 2022년도 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2021R1A2C3013687).

참고문헌

[1] Q. Yang, Y. Liu, T. Chen, and Y. Tong, “Federated machine learning: Concept and applications,” ACM Trans. Intell. Syst. Technol., vol. 10, no. 2, pp. 1–19, 2019.
[2] E. B. P. Kairouz and H. B. Mcmahan, “Advances and open problems in federated learning,” Found. Trends Mach. Learn., vol. 14, no. 1, pp. 1–210, 2021.
[3] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and
A. T. Suresh, “SCAFFOLD: Stochastic controlled averaging for federated
learning,” in Proc. ICML, 2020, pp. 5132–5143.
[4] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, “On the convergence of FedAvg on non-IID data,” in Proc. ICLR, 2020, pp. 1–26.
[5] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Proc. AISTATS, 2017, pp. 1273–1282.
[6] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” in Proc. Mach. Learn. Syst., vol. 2, 2020, pp. 429–450.
[7] H. Wang, Z. Kaplan, D. Niu, and B. Li, “Optimizing federated learning on non-IID data with reinforcement learning,” in Proc. IEEE Conf. Comput. Commun. (INFOCOM), Jul. 2020, pp. 1698–1707.
[8] M. Yang, X. Wang, H. Zhu, H. Wang, and H. Qian, “Federated learning with class imbalance reduction,” in Proc. 29th Eur. Signal Process. Conf. (EUSIPCO), Aug. 2021, pp. 2174–2178.
[9] LeCun, Yann, and Yoshua Bengio. "Convolutional networks for images, speech, and time series." The handbook of brain theory and neural networks 3361.10 (1995): 1995.
[10] Xiao, Han, Kashif Rasul, and Roland Vollgraf. "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms." arXiv preprint arXiv:1708.07747 (2017).