연합학습에서의 사용자 경험 기반 앙상블 기법

김형빈, 박현희*
명지대학교 정보통신공학과

hbkim@mju.ac.kr, hhpark@mju.ac.kr*

User Experience-based Ensemble Method in Federated Learning

Hyungbin Kim and Hyunhee Park*
Department of Information and Communication Engineering, Myongji University

요 약

본 논문은 연합학습 모델의 정확도 및 손실 함수 측면에서 성능 향상을 위한 사용자 경험 기반 앙상블 기법을 제안한다. 제안된 기법은 각 무선 기기의 심층 신경망 학습 결과를 서버에 보내는 데에만 국한하지 않고, 이후 연합학습의 진행 단계에서 각 무선 기기의 앙상블이 이루어지도록 활용한다. 각 무선 기기는 자신의 데이터를 통해 이전에 학습된 심층 신경망을 앙상블에 활용함으로써 사용자 경험 기반의 연합학습 모델을 얻을 수 있다. 모의실험을 통해, 제안된 사용자 경험 기반 앙상블 기법이 적용된 연합학습 모델이 기존 연합학습 모델에 비해 손실 함수 측면에서 13% 향상된 성능을 보임을 확인할 수 있다.

Ⅰ. 서 론

연합학습은 다수의 무선 기기와 하나의 서버로 구성된 분산형 기계 학습 기술이다 [1]. 각 무선 기기는 자신이 갖고 있는 데이터를 바탕으로 장치 내부에서 심층 신경망 학습을 진행하며, 심층 신경망의 학습 결과를 서버에 전송하는 구조를 갖는다. 서버는 각 무선 기기로부터 전송된 학습 결과를 취합하여 전역 심층 신경망을 학습한다. 이와 같이 각 무선 기기의 데이터를 서버에 전송하지 않아도 전역 심층 신경망의 학습을 진행할 수 있는 연합학습의 특성으로 인해, 기존의 심층 신경망 기법들과 달리 데이터의 익명성이 보장되는 특징을 갖는다. 그러나 실제 환경에서 각 무선 기기가 갖는 데이터가 독립항등분포(Independent and Identically Distributed)를 보일 확률은 매우 작으며 분산이 클수록 학습된 전역 심층 신경망의 성능이 저하되는 문제가 발생한다 [2].

데이터가 독립항등분포를 보이지 않을 때 발생하는 문제를 해결하기 위한 연구들을 살펴보면, 무선 기기와 서버 간의 심층 신경망 전송이 이루어질 때 분산 개념을 제안하며 기존 연합학습에 비해 향상된 성능을 보였다 [3]. 하지만 해당 연구는 무선 기기와 서버간 통신이 이루어질 때 학습 결과와 함께 분산 값이 전송되어야 하기 때문에 기존 연합학습보다 2 배 큰 통신비용이 발생하는 한계점이 있다. 각 무선 기기에서 생성된 심층 신경망을 서버에서 취합하고 표준편차를 계산하여, 평균에서 표준편차 범위 내의 값을 갖는 무선 기기만을 이후 학습에 참여하도록 하는 연구가 진행되었다 [4]. 해당 연구에서 제안하는 FedSD 를 통해 왜곡된 학습을 유발하는 무선 기기를 제외하고 학습을 진행하여, 정확도 및 손실 함수 측면에서 기존 연합학습에 비해 향상된 성능을 보인다. 그러나 이 경우 연합학습에 전혀 참여하지 못하는 무선 기기가 발생하게 되는 한계점이 있다.

본 논문에서는 연합학습 내 무선 기기들의 데이터가 독립항등분포를 갖지 않는 상황에서 연합학습 모델의 성능을 향상시킬 수 있는 사용자 경험 기반 앙상블 기법을 제안한다. 제안된 기법의 핵심 내용은 지역 심층 신경망의 학습 과정에 있다. 무선 기기가 이전 단계에서 자신의 데이터로 학습된 지역 심층 신경망과 서버로부터 전송된 전역 심층 신경망을 앙상블하고, 앙상블 후 생성되는 심층 신경망을 기반으로 다음 단계의 지역 심층 신경망을 학습하는 기법을 제안한다.

Ⅱ. 본론

본 논문에서는 연합학습에 참여하는 무선 기기들이 동일한 형태의 심층 신경망을 사용한다고 가정한다. $local\ model_n$은 그림 1 의 $(t-1)$ 단계에서 생성된 각 지역 심층 신경망이며, n 은 무선 기기의 번호를 나타낸다. $global\ model_t$ 는 모든 n에 대한 $local\ model_n$ 이 서버로 전송된 이후 (t) 단계에서 생성된 전역 심층 신경망이다. 전역 심층 신경망은 취합된 모든 지역 심층 신경망을 일반화한 형태이다. 때문에 (t) 단계에서 $global\ model_t$ 를 기반으로 각 무선 기기의 $local\ model_n$을 학습하게 되면, $(t-1)$ 단계의 $local\ model_n$으로 이어서 학습하는 것보다 성능이 저하될 수 있다.

기존 연합학습은 (t) 단계에서 각 무선 기기의 지역 심층 신경망을 학습하기 위해 $global\ model_t$ 를 초기 모델로 사용한다. 본 논문에서 제안된 기법은 $global\ model_t$가 갖는 일반화의 오류를 낮추고, 무선 기기의 데이터로 표현되는 사용자 경험을 기반으로 (t) 단계가 진행되도록 한다. 이를 위해 $global\ model_t$ 와 $(t-1)$ 단계에서 생성된

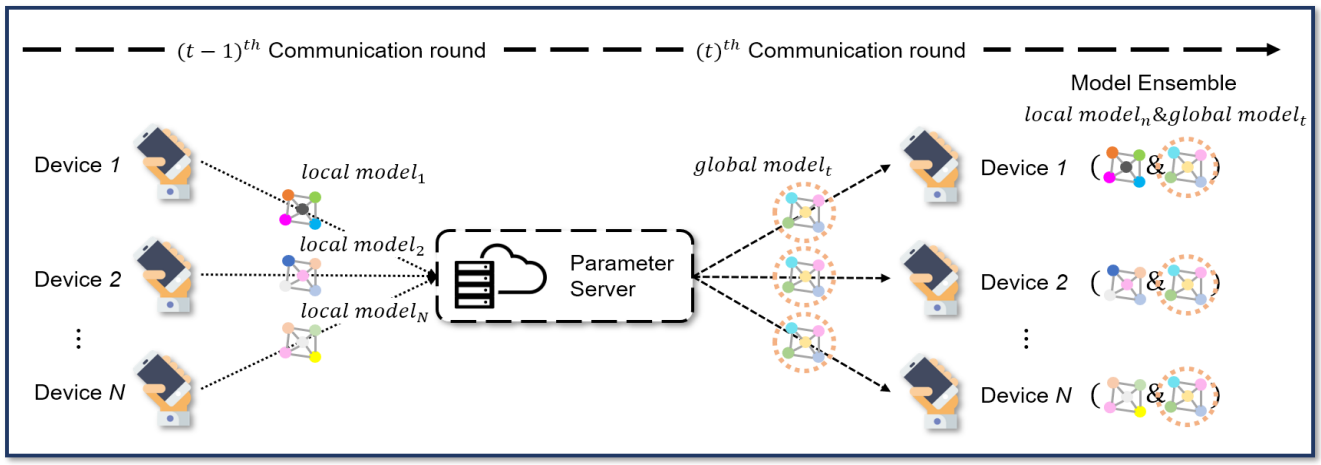

그림 1. 제안된 기법의 구조

알고리즘 1: 제안된 기법의 알고리즘

Input: local datasets $\mathcal{D}^n$, number of devices N, number of local epoch T, learning rate η, global model $global\ model_t$, $(t-1)^{th}$ local model $local\ model_n$

Output: $(t)^{th}$ local model $new\ local\ model_n$

LocalTraining:

$new\ local\ model_n \leftarrow \frac{local\ model_n + global\ model_t}{2}$

for device n = 1, 2, …, N **in parallel do**

  for local epoch t = 1, 2, …, T **do**

    for each batch $\boldsymbol{b}$ = {$\mathbf{x}$, y} of $\mathcal{D}^n$ **do**

      $new\ local\ model_n \leftarrow \eta \nabla \ell(new\ local\ model_n; b)$

return $new\ local\ model_n$ to server

$local\ model_n$ 을 앙상블하여 초기 모델로 사용한다. 이 과정을 의사코드로 작성하면 알고리즘 1 과 같다.

본 논문에서 제안된 사용자 경험 기반 앙상블 기법을 활용한 연합학습 모델의 성능을 평가하기 위해 모의실험을 진행하였다. 심층 신경망은 LeNet-5 모델을 사용하였으며, MNIST 데이터셋을 사용하여 학습을 진행하였다 [5]. 연합학습에 참여하는 무선 기기는 10 개, 학습 반복 횟수는 100 회로 설정하였다. 이때 각 무선 기기가 갖는 데이터 개수와 클래스를 다르게 할당하여 독립항등분포를 보이지 않도록 하였다. 동일한 실험 환경을 갖는 기존 연합학습들을 진행하여 본 논문에서 제안된 연합학습과 성능을 비교한다.

그림 2 는 모의실험의 결과를 나타낸다. 아래 결과는 본 논문에서 제안된 사용자 경험 기반 앙상블 기법을 적용하였을 때, 데이터 분포가 독립항등분포를 보이지 않는 상황에서 연합학습 모델의 성능이 향상될 수 있음을 보인다.

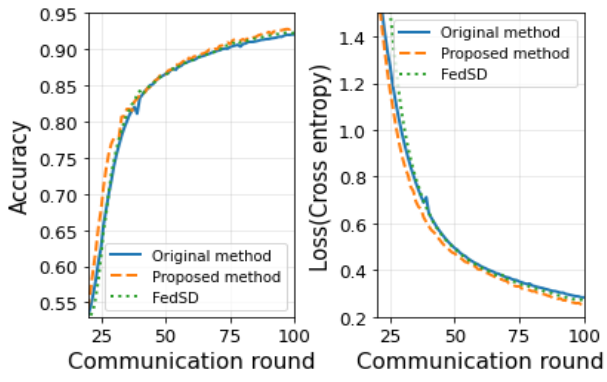

그림 2. 제안된 기법과 기존 기법의 성능

Ⅲ. 결론

본 논문에서는 연합학습에서 무선 기기의 데이터 분포가 독립항등분포를 보이지 않을 때 연합학습 모델의 성능이 저하되는 문제 개선을 목표로 연구를 진행하였다. 본 논문에서는 연합학습의 지역 심층 신경망 학습 단계에서 사용자 경험 기반 앙상블 기법을 활용한 연합학습 모델을 제안하였다. 제안된 기법은 학습 과정에서 무선 기기의 데이터가 서버로 전송되지 않기 때문에 데이터의 익명성을 보장하며, 모의실험을 통해 제안된 기법을 적용한 연합학습 모델이 기존 연합학습 모델보다 성능이 향상됨을 보였다. 향후 연구에서는 탈중앙화 연합학습의 지역 심층 신경망 학습 단계에서 앙상블 기법을 적용하는 방식에 대한 연구를 진행할 예정이다.

ACKNOWLEDGMENT

이 논문은 2021 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00368, 6G 서비스를 위한 인공지능/머신러닝 기반 자율형 MAC 개발, No. 2021-0-00990, 설명가능한 인공지능 기반 무선랜 네트워크 시스템 고도화 핵심 기술 연구).

참 고 문 헌

[1] J. Konecny, H. B. McMahan, F. X. Yu, P. Richtarik, A. T. Suresh, and D. Bacon, “Federated learning: Strategies for improving communication efficiency.” arXiv:1610.05492v2 [cs.LG], Oct. 2017.

[2] Y. Zhao, M. Li, L. Lai, N. Suda, D. Civin, and V. Chandra, “Federated learning with Non-IID data.” arXiv:1806.00582 [cs.LG], Jun. 2018.

[3] S. P. Karimireddy, S. Kale, M. Mohri, S. J. Reddi, S. U. Stich, and A. T. Suresh, “Scaffold: Stochastic controlled averaging for on-device federated learning.” arXiv:1910.06378 [cs.LG], Oct. 2019.

[4] H. Kim, Y. Kim, G. Woo, J. Kim, and H. Park, "FedSD: Federated Learning algorithm with Standard Deviation of weights for each user." Proceedings of Symposium of the Korean Institute of communications and Information Sciences, pp. 662-663, June. 2021.

[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition.” Proceedings of the IEEE 86.11, pp. 2278- 2324, Nov. 1998.