

초저지연 환경 원샷 분산 특징 학습 기반 데이터 중요도 인지 스케줄링 통합 프레임워크 설계

Design of a Unified Framework for Data-Importance-Aware Scheduling Based on One-Shot Distributed Feature Learning in Ultra-Low Latency Environments

Ye-Jin Kim, Won-Yong Shin

Yonsei University

ay0119@yonsei.ac.kr, wy.shin@yonsei.ac.kr

요 약

본 논문은 분산 학습 시스템에서 원샷으로 특징 학습 (feature learning)을 수행하고 데이터 중요도 (DI: data importance) 인지 스케줄링을 수행하는 통합 프레임워크를 제안한다. 구체적으로, 원샷 분산 특징 학습을 위해 저차원 표현 학습 과정 중 각 디바이스와 서버는 단 한 라운드의 통신으로 디바이스 간 저차원 표현 벡터들이 서로 정렬되도록 하는 최대 평균 불일치 (MMD: maximum mean discrepancy) 손실을 포함한 새로운 손실 함수를 설계한다. 또한, 데이터 중요도 인지 스케줄링을 수행하기 위해 전체 데이터셋과 지역 데이터셋의 분포 간 유사도를 정량화하여 표현할 수 있는 새로운 성능 지표인 DI 를 소개하고, DI 최대화 측면에서 디바이스를 선택하는 방법을 설명한다. 제안하는 통합 프레임워크는 초저지연 환경에 적합하기에 기존 연합 학습 방법이 적용되기 어려운 분산 학습 시스템에서 기계 학습 문제를 해결하는데 효과적으로 사용될 수 있다.

I. 서 론

연합 학습 (federated learning)은 분산 학습 시스템에서 디바이스의 모델 파라미터를 온-디바이스 학습을 통해 반복적으로 엣지 서버와 공유하여 모든 디바이스들의 로컬 모델 파라미터를 동기화하는 방식이다. 하지만 지연에 민감한 초저지연 (ultra-low latency) 시스템에서는 반복적인 모델 파라미터 업데이트가 제한적이기에 이를 해결하기 위한 다양한 챌린지가 존재한다. [1]에서는 디바이스 간 저차원 표현 벡터들이 서로 정렬되어 있다는 이상적인 가정 하에 분산된 지역 데이터의 다양성 (diversity)을 최대한 활용하는 측면에서 디바이스를 선택하는 새로운 데이터 중요도 (DI: data importance) 인지 스케줄링을 제안하였다. [2]에서는 [1]에서의 학습 시 사용한 가정을 완화할 수 있는 해결책으로써, 서버와 디바이스 간 단 한 라운드의 통신으로 디바이스 별 데이터의 저차원 표현 벡터들이 서로 정렬되도록 하는 최대 평균 불일치 (MMD: maximum mean discrepancy) 손실을 포함한 새로운 손실 함수를 설계하고 이를 바탕으로 원샷 분산 특징을 학습하는 프레임워크를 제안하였다. 본 논문에서는 데이터가 다수의 디바이스에 분산되어 있고 엣지 서버가 전체 데이터 셋에 대해 소량의 정보만 가지고 있는 분산 학습 시스템을 고려하고, [1]과 [2]에서 제안한 방법을 통합하여 원샷으로 특징 학습을 수행하고 데이터 중요도 인지 스케줄링을 수행하는 통합 프레임워크를 제안한다.

II. 본론

첫째, 원샷 분산 특징 학습 프레임워크에 대해 설명한다. 저차원 표현 벡터 정렬을 위해 서버는 소량의 정보로 학습한 얇은 모델 파라미터를 디바이스에게 브로드캐스팅한다. 각 디바이스는 이를 임베딩 추출을 위한 모델 학습에 사용한다. 임베딩 공간에서의 디바이스

별 데이터의 임베딩과 서버에서 얇은 모델 학습용 데이터의 임베딩에 대해 커널 함수를 이용하여 두 임베딩 간 차를 MMD 로 정의한다. 이 때, MMD 와 cross-entropy 와 같은 분류 손실 함수를 결합한 새로운 손실 함수를 설계하고, 이에 기반하여 각 디바이스에서 모델을 학습한 뒤 취득한 임베딩의 분포를 가우시안 혼합 모델 (Gaussian mixture model)을 사용하여 추정하여 GMM 파라미터들을 서버에게 상향링크를 통해 전송한다.

둘째, 데이터 중요도 인지 스케줄링에 대해 설명한다. 서버에서는 디바이스로부터 전송 받은 GMM 파라미터들을 활용하여 DI 기반 최대 우도 (maximum likelihood) 측면에서 디바이스 집합을 선택하는 스케줄링을 수행한다. 이 때, 한 번에 전체 디바이스 집합을 선택하는 것은 NP-hard 하므로, greedy 알고리즘을 사용하여 순차적으로 디바이스를 선택한다.

ACKNOWLEDGMENT

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1A2C3004345) and by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2021-0-00347, 6G Post-MAC (POsitioning- & Spectrum-aware intelligent MAC for Computing & Communication Convergence)).

참 고 문 헌

- [1] J.-Y. Kim, Y.-J. Kim, and W.-Y. Shin "Data importance-aware scheduling for exploiting diversity of distributed data," in *Proc. KICS-Winter Conference*, Feb. 2022.
- [2] Y.-J. Kim, J.-Y. Kim, and W.-Y. Shin, "A framework for one-shot distributed feature learning," in *Proc. KICS-Summer Conference*, Jun. 2022.