

이질적 데이터 환경에서의 개인화 성능 향상을 위한 프로토타입 기반 연합 프롬프트 학습

박지원, 이주형*
가천대학교

pich7755@gachon.ac.kr, *j17.lee@gachon.ac.kr

Prototype-based Federated Prompt Learning for Personalization in Heterogeneous Environments

Jiwon Park, Joohyung Lee*

Department of Computing at Gachon University

요약

최근 파운데이션 모델의 발전과 함께, 연합 학습 환경에서 사전 학습된 모델의 프롬프트만을 학습하는 Federated Prompt Learning (FPL) 방식이 통신 효율성 측면에서 주목받고 있다. 그러나 실제 연합 학습 환경에서는 클라이언트마다 데이터 분포가 서로 다른 통계적 이질성이 존재하며, 이는 단순한 프롬프트 파라미터의 집계만으로는 개별 클라이언트의 특성에 맞는 최적의 개인화 성능을 보장하기 어렵다는 한계를 가진다. 이를 해결하기 위해 본 논문에서는 모델의 출력 임베딩을 클래스별로 요약한 추상적 프로토타입을 교환하는 새로운 학습 프레임워크를 제안한다. 제안 기법은 프로토타입의 교환 및 정규화를 통해, 파라미터 공유 없이도 이질적 클라이언트 간 효과적인 지식 전이를 수행한다. 이를 통해 FPL의 통신 효율성을 유지하면서 모델의 수렴 속도와 개인화 정확도를 획기적으로 향상시킨다.

I. 서론

통신 대역폭과 컴퓨팅 자원이 제한적인 모바일·엣지 환경에서 연합 학습(Federated Learning, FL)의 실현 가능성은 통신 효율성에 크게 의존한다. 최근 대규모 파운데이션 모델의 발전과 함께, 전체 모델 파라미터 대신 학습 가능한 프롬프트만을 업데이트 및 교환하는 Federated Prompt Learning(FPL)이 통신 비용을 획기적으로 절감하는 대안으로 주목받고 있다. 이 방식은 거대 모델의 백본 가중치를 고정된 상태에서 극소수의 파라미터만을 전송하므로, 모델 전체를 공유해야 했던 기존 방식 대비 네트워크 대역폭 소모를 수천 배 이상 줄일 수 있다는 장점이 있다[1].

그러나 기존 FPL은 통계적 이질성 환경에서 개인화 성능에 한계를 보인다. 클라이언트들은 서로 다른 레이블 분포(Non-IID)를 가지며, 중앙 서버에서의 단순한 프롬프트 평균화는 로컬 데이터 특성을 충분히 반영하지 못한다. 이로 인해 글로벌 프롬프트는 평균적인 표현에 머물거나 특정 클라이언트에 편향되어 개인화 성능과 학습 수렴 효율을 저하시킨다[2].

이에 본 논문에서는 FPL의 통신 효율성을 유지하면서 데이터 이질성 환경에서 개인화 성능을 향상시키기 위해, 프롬프트 집계와 프로토타입 집계를 결합한 FPL 프레임워크를 제안한다[3]. 프롬프트 파라미터는 기존과 같이 집계하고, 각 클라이언트는 클래스별 임베딩 평균인 클래스 프로토타입을 함께 전송한다. 서버는 글로벌 프로토타입을 생성·배포하며, 클라이언트는 이를 로컬 학습의 정규화 신호로 활용하여 데이터 이질성 환경에서도 강건한 개인화 모델을 학습한다. 이는 단순한 파라미터 평균화의 한계를 넘어 의미적 차원의 지식 공유를 가능케 함으로써, 클라이언트 간의 정보 불일치를 해소하고 학습 효율성을 제고하는 데 기여한다.

II. 본론

그림 1은 제안하는 프로토타입 기반 FPL 프레임워크의 전체 구조를 나타낸다. 각 클라이언트는 인코더를 고정된 채 프롬프트와 MLP만을 학습하며, 클래스별 로컬 프로토타입을 생성해 서버로 전송한다. 서버는 수집된 프롬프트와 프로토타입을 집계하여 다시 배포하며, 클라이언트는 이를 활용해 손실 함수를 최소화하는 방향으로 학습을 진행한다. 이러한 구조는 데이터 이질성 환경에서도 효과적인 지식 공유를 통해 개인화 성능을 극대화하도록 설계되었다.

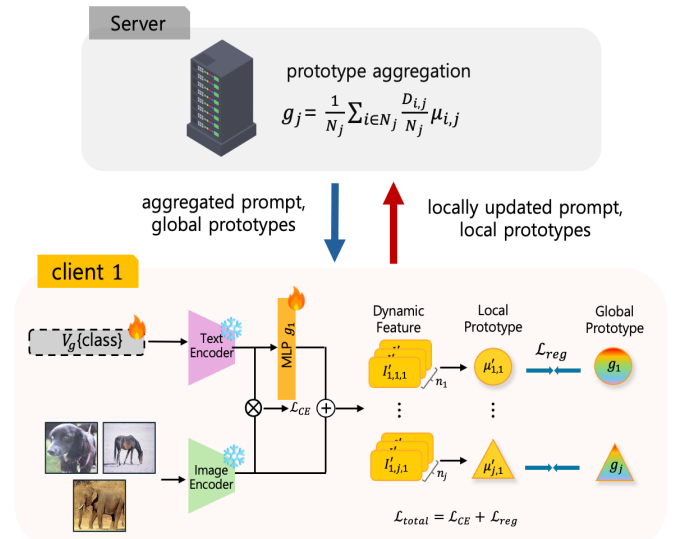


그림 1. 제안하는 프로토타입 기반 FPL 프레임워크의 전체 아키텍처.

# methods	Personalized test acc		
	Oxford Pets	Caltech101	Avg .
FedPGP	87.54	90.8	89.17
FedOTP	88.21	91.17	89.69
Ours	89.78	91.36	90.57

표 1. 이질적인 데이터셋 환경에서의 개인화 테스트 정확도 비교 결과.

가. 프롬프트 학습

파운데이션 모델을 다운스트림 작업에 효과적으로 적용하기 위한 파라미터 효율적 미세조정 기법으로서 프롬프트 학습이 널리 활용되고 있다. 이 방식은 대규모 사전 학습 모델의 파라미터 대부분을 동결한 상태로 유지하고, 입력 임베딩 공간에 소수의 학습 가능한 연속 벡터만을 추가하여 최적화를 수행한다. 모델의 모든 가중치를 업데이트하는 기존의 전체 미세조정과 비교했을 때, 프롬프트 학습은 훈련이 필요한 파라미터의 수를 극적으로 줄여주는 이점이 있다. 이에 따라 학습에 소요되는 시간과 메모리 요구량이 현저히 감소하므로, 컴퓨팅 자원이 제한적인 환경에서도 거대 모델의 운용 효율성을 크게 높일 수 있다.

나. 프로토타입 기반 정규화

데이터 이질성 환경에서 단순 프롬프트 집계는 클라이언트의 고유한 분포 특성을 반영하지 못해 성능 저하를 야기한다. 이를 해결하기 위해 본 논문은 클래스별 임베딩 벡터의 평균인 프로토타입을 도입한다. 클라이언트가 생성한 로컬 프로토타입을 서버가 집계하여 글로벌 프로토타입을 생성하는 이 과정은, 구체적인 파라미터 교환 없이도 클라이언트 간의 의미적 지식을 효율적으로 공유할 수 있게 한다.

$$\mathcal{L}(D_i, \omega_i) = \mathcal{L}_{CE}(\mathcal{F}_i(\omega_i; x), y) + \mathcal{L}_{reg}(\mu_i, g) \quad (1)$$

여기서, \mathcal{L}_{CE} 는 분류 손실을 μ_i 와 g 는 각각 로컬 및 글로벌 프로토타입을 나타낸다. 식 (1)의 정규화 항(\mathcal{L}_{reg})은 서버로부터 수신한 글로벌 프로토타입(g)을 기준으로 로컬 프로토타입(μ_i)이 정렬되도록 유도한다. 이는 모델이 편향된 로컬 데이터에 과적합되는 것을 방지하고 특징 공간의 일관성을 확보하는 핵심적인 가이드 역할을 수행한다. 결과적으로 식 (1)에 기반한 학습은 추가적인 통신 비용 없이도 이질적인 환경에서 모델의 수렴 안정성과 개인화 성능을 효과적으로 강화한다.

다. 성능 분석

본 절에서는 제안하는 프로토타입 기반 FPL 프레임워크의 성능 우수성을 입증하기 위해, 대표적인 이미지 분류 데이터셋인 Oxford Pets와 Caltech101에서의 실험 결과를 비교 분석한다. 비교 대상으로는 최신 FPL 방법론인 FedPGP와 FedOTP를 선정하였으며, 각 클라이언트의 로컬 데이터 분포에 최적화된 성능을 평가하기 위해 개인화 테스트 정확도를 주요 지표로 삼았다.

표 1에 제시된 바와 같이, 제안 기법은 두 데이터셋 모두에서 기존 베이스라인 모델들을 상회하는 정확도를 달성하였으며, 이는 프로토타입을 활용한 의미적 지식 공유가 데이터 이질성 환경에서 개인화 성능 향상에 유의미하게 기여함을 시사한다. 특히, 단순한 프롬프트 공유에 그치는 기존 방식들과 달리, 프로토타입을 통해

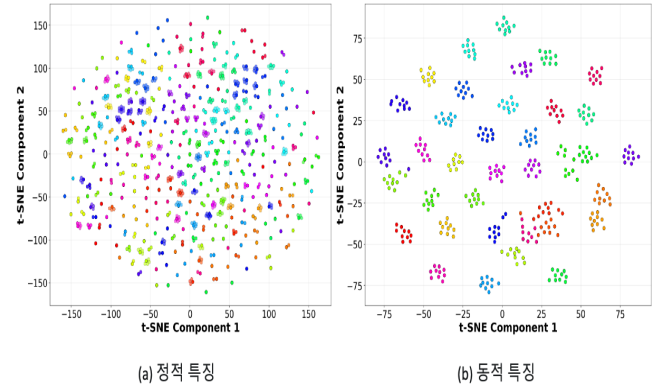


그림 2. t-SNE를 이용한 특징 공간 시각화 비교.

특징 공간을 직접적으로 교정함으로써 모델 간의 불일치를 최소화했다는 점에서 기술적 의의가 크다.

그림 2는 학습 과정에서 추출된 정적 특징(a)과 제안 기법 적용 후의 동적 특징(b) 분포 변화를 t-SNE로 시각화하여 비교한다. 초기 정적 특징 공간(a)에서는 다양한 클래스의 샘플이 혼재되어 경계가 모호한 반면, 프로토타입 기반 학습이 적용된 동적 특징 공간(b)에서는 동일 클래스 샘플들이 강하게 응집하며 클래스 간 명확한 분리가 나타난다. 이는 제안하는 프로토타입 정규화가 이질적인 환경에서도 특징 표현을 효과적으로 교정하여 분류 성능을 향상시킴을 시각적으로 입증한다.

III. 결론

본 논문에서는 FPL 환경의 데이터 이질성 한계를 극복하고 개인화 성능을 극대화하기 위해, 프로토타입을 활용하여 클라이언트 간의 지식을 정렬하는 새로운 프레임워크를 제안하였다. 실험 결과, 제안 기법은 기존 베이스라인 모델 대비 우수한 개인화 정확도를 달성하였으며, t-SNE 분석을 통해 특징 공간에서의 명확한 클래스 군집화 효과를 입증하였다. 이는 추가적인 파라미터 공유 없이 프로토타입 기반의 정규화만으로도 통신 효율성을 유지하면서 로컬 모델의 성능을 효과적으로 개선할 수 있음을 시사한다. 향후 연구에서는 프로토타입 전송 과정에 경량화된 프라이버시 보호 기술을 접목하여, 제안 구조의 보안성과 실용성을 더욱 높이는 방향으로 연구를 확장할 계획이다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 2026년도 SW 중심대학사업의 결과로 수행되었음(2021-0-01389)

참고 문헌

- [1] Guo, Tao, et al. "Promptfl: Let federated participants cooperatively learn prompts instead of models-federated learning in age of foundation model." *IEEE Transactions on Mobile Computing* 23.5 (2023): 5179-5194.
- [2] Tan, Alysa Ziyang, et al. "Towards personalized federated learning." *IEEE transactions on neural networks and learning systems* 34.12 (2022): 9587-9603.
- [3] Tan, Yue, et al. "Fedproto: Federated prototype learning across heterogeneous clients." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 36. No. 8. 2022.