

뇌연령 예측을 위한 연합학습 모델 개발*

문성환¹ 한주혁 이원희

경희대학교 소프트웨어융합학과

msh2044@khu.ac.kr, hhannn@khu.ac.kr, whlee@khu.ac.kr

Federated Learning for Brain Age Prediction

Sunghwan Moon Juhyuk Han Won Hee Lee

Department of Software Convergence, Kyung Hee University

요 약

연합학습(Federated Learning, FL)은 다수의 로컬 클라이언트와 하나의 중앙 서버가 협력하여 분산된 훈련 데이터를 사용하여 글로벌 모델을 훈련시킬 수 있는 새로운 학습 패러다임이다. 로컬 클라이언트의 데이터 전송 없이 학습한 모델의 가중치를 계산하고 이를 중앙 서버에서 취합하여 전체 모델을 향상시킬 수 있다. 개인 정보의 수집과 활용에 대한 규제가 강화됨에 따라 직접적인 데이터 공유가 필요없는 연합학습의 적용이 의료 분야로 확대되고 있다. 본 연구에서는 자기공명영상(Magnetic Resonance Imaging, MRI) 데이터를 이용한 기계학습 기반 뇌연령 예측을 위한 연합학습 모델을 개발하였다. 중앙 서버는 학습을 위한 로컬 기계학습 기반 뇌연령 예측 모델을 각 클라이언트에게 전송하고, 사전에 클라이언트에게 승인받은 데이터 저장공간에 대한 포인터를 이용하여 로컬 환경에서 각 클라이언트들이 뇌연령 예측 모델을 학습시킬 수 있도록 한다. 이후 중앙 서버는 클라이언트에서 학습한 모델의 가중치를 수신하여 평균화하고 다시 클라이언트에게 전송하여 글로벌 뇌연령 예측 모델을 향상시킨다. FL을 적용하지 않은 non-FL 모델과 비교했을 때 FL 모델을 통해 뇌연령 예측을 한 결과 Mean Absolute Error(MAE)는 2.7-31.7% 낮아졌으며, 상관관계(Pearson correlation coefficient, R)값은 최대 7.4% 향상되었다. 본 연구에서 개발한 연합학습을 이용하여 환자의 개인정보 공유 없이 보다 정확한 뇌연령 예측이 가능해졌으며 신경퇴행성 및 정신질환 환자의 가속 노화 예측력을 향상시킬 수 있을 것으로 기대한다.

1. 서 론

인공지능(Artificial Intelligence, AI)의 발전에 따라, 건강데이터(Lifelog Data), 전자의무기록(Electronic Medical Record, EMR), 자기공명영상(Magnetic Resonance Imaging, MRI) 영상 등 의료 데이터를 활용한 연구 및 서비스가 확장되고 있다. 하지만 최근 개인정보의 수집과 활용에 대한 규제가 강화되면서, 의료데이터를 활용한 연구가 제약을 받고 있다. 유럽 연합의 경우 2018년부터 '일반 데이터 보호 규칙(General Data Protection Regulation, GDPR)'을 통해 개인정보 보호에 관한 규제를 강화하였으며, 우리나라의 경우 2020년 '데이터 3법'이 국회 본회의의 통과하면서 개인정보 활용 기준이 엄격해졌다[1].

의료 데이터 등과 같이 개인정보를 포함하는 데이터의 활용이 어려워지는 문제를 해결하기 위한 방안으로 최근 연합학습(Federated Learning, FL)이라는 새로운 학습 패러다임이 소개되었다. FL은 사용자의 데이터를 직접적으로 공유하지 않고, 사용자의 로컬 환경에서 학습한 모델의 가중치 공유를 통해 전체 모델을 향상시키는 새로운 학습 기법이다. FL에서는 각 클라이언트들이 각자의 로컬 데이터를 통해 학습하여 나온 모델의 가중치를 중앙 서버로 송신하고, 중앙 서버는 클라이언트로부터 수신 받은 가중치를 취합하여 글로벌 가중치를 생성 후 클라이언트에게 배포한다[2]. FL은 직접적인 데이터 공유 없이 학습된 모델의 가중치만을 중앙 서버로 전송하기 때문에 개인정보를 보호할 수 있을 뿐만 아니라 암호화 작업도 수월하게 진행할 수 있다.

최근, 엔비디아(NVIDIA)는 킹스 칼리지 런던(King's College London)과 협업하여 뇌종양(Brain Tumor) 분할(segmentation)을 위한 FL 모델의 성능을 테스트한 결과를 보고했다. 표본 242명의 뇌종양 MRI 데이터를 이용하여 로컬 연합 클라이언트 13곳으로 데이터를 분배 후 FL을 진행하였다. 뇌종양 분할을 위한 연합학습 모델 결과가 중앙집중형(Data-centralized) 학습 모델 결과와 유사한 성능을 보여주었다[3].

본 연구에서는 T1 강조 자기공명영상 데이터를 이용하여 뇌영역별 형태학적(morphological) 지표를 계산하여 로컬 기계학습 모델의 feature로 활용하였다. 뇌연령 예측을 위한 로컬 기계학습 모델을 위해 다층 퍼셉트론(Multi-Layer Perceptron, MLP) 방법을 사용하였다. 예측된 뇌연령과 실제 연령의 차이(뇌연령 차이, brain-age gap)는 정신질환 및 신경퇴행성 질환과 관련되어 있어 의료 분야에서 새로운 생체지표로 제시되어 연구가 진행되고 있다[4]. 하지만 개인정보 보호로 인해 많은 의료 데이터를 활용할 수 없어 고성능의 뇌연령 예측모델을 개발하는데 제약을 받고 있다. 따라서 본 연구에서 로컬 기계학습 기반 뇌연령 예측 모델을 연합학습 모델에 적용하여 성능을 비교 및 평가하였다. 본 연구에서 개발된 연합학습 기반 뇌연령 예측 모델을 통해 개인 데이터 보호와 함께 향상된 뇌연령 예측 정확도를 보여주고자 한다.

* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심 대학지원사업의 연구결과로 수행되었음(2017-0-00093)

2. 연구 방법

2.1 데이터 수집

자기공명영상 데이터를 수집하기 위해 온라인 공유 뇌영상 데이터베이스를 이용하였다. Cambridge Centre for Ageing and Neuroscience(Cam-CAN) 데이터베이스를 이용하여 T1 강조 자기공명영상 데이터(표1)를 수집하였다[5].

표 1. Cam-CAN 데이터 수 및 연령

Data Set	데이터 수	최소 연령	최대 연령	평균 연령	표준 편차
Cam-CAN	601	18	88	54.40	18.13

2.2 뇌영역 형태학적 Feature 추출

Freesurfer(버전 v7.1.0)[6]의 recon-all 파이프라인을 사용하여, T1 강조 자기공명영상을 전처리하고 Desikan-Killiany atlas와 probabilistic atlas를 이용하여 세분화된 뇌영역별 형태학적 지표들을 추출하였다. 각 subject의 피질 두께(cortical thickness, n=68), 피질 표면 넓이(cortical surface area, n=68), 피질하 부피(subcortical volume, n=16), 그리고 total intracranial volume를 계산하여 기계학습 모델의 feature로 사용하였다.

2.3 연합학습 프레임워크

연합학습 프레임워크는 실제 네트워크 연결을 통해 FL을 구현할 수 있는 PySyft 라이브러리를 이용하여 개발하였다[7]. 연합학습의 프레임워크 구성을 위해 데이터셋은 2곳의 연합 클라이언트에게, (1) 정규 분포 형태와 (2) 편향 분포 형태, 두가지 로컬 환경을 만들어 연합학습의 성능을 평가하였다. FL 모델에서는 각 로컬 클라이언트에 배포된 모델(다층 퍼셉트론)을 사용하여 로컬별로 학습을 진행했고, 최종 학습된 파라미터들을 중앙 서버에 전송하였다. 중앙 서버는 각 로컬에서 송신한 가중치를 취합하여 평균 가중치(글로벌 가중치)를 계산하고 다시 로컬 클라이언트에게 배포했다. 연합학습 모델을 사용하지 않는 로컬 모델(이하 non-FL)은 FL에서 배포된 글로벌 가중치를 사용하지 않고 독립적으로 학습을 진행하여 성능을 평가했다. FL모델과 non-FL 모델을 통해 예측되어진 뇌연령과 실제 연령의 Mean Absolute Error(MAE)값과 상관관계(Pearson correlation coefficient, R)값을 계산하여 성능을 비교 및 평가하였다.

2.3.1 데이터 분포

기계학습 알고리즘 적용전 scikit-learn[8] 라이브러리에서 제공하는 Standard Scaler를 사용하여 feature값을 정규화하였다. 뇌영역별 형태학적 feature 데이터셋(n=601)중 500개는 학습용 데이터로 101개는 테스트 데이터셋으로 사용하였다.

전처리된 데이터는 클라이언트 2곳으로 각각 학습용 데이터 250개씩 분배하였다. 데이터는 크게 두가지 경우로 나누어 분배되었다. 정규분포 형태의 데이터셋을 분배한 경우와 편향분포 형태의 데이터셋으로 나누어 연구를 진행했다.

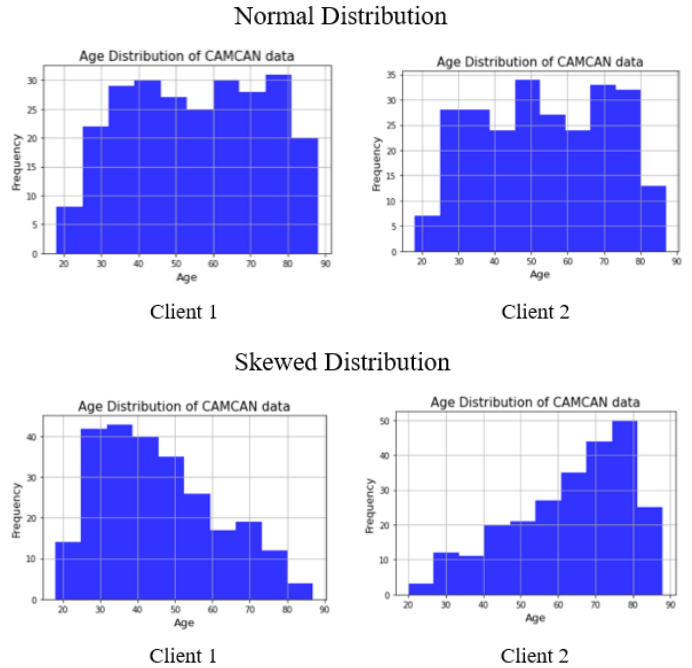


그림 1. 데이터 분포

2.3.2 연합학습 아키텍처

PySyft 기반 연합학습 환경에서 각 클라이언트들은 AWS에 로컬 데이터를 저장하고, 중앙 서버에게 데이터 포인터에 대한 접근 권한을 부여할 수 있다. 중앙 서버는 각 클라이언트에게 로컬 학습 모델(MLP), 학습률, 최적화 방법 등을 배포하며, 사전에 허가된 데이터 포인터를 제어하여 배포된 모델에 데이터셋을 삽입한다. 배포된 모델과 데이터셋은 로컬 환경에서 훈련되며, 훈련이 종료된 이후에는 중앙서버에게 훈련된 가중치를 전송한다. 중앙서버는 전송된 가중치들을 평균화(Federated Averaging)하여 클라이언트에게 다시 전송하고 1 라운드가 종료된다[9]. 본 연구에서는 총 3 라운드를 사용하여 실험하였다(그림 2).

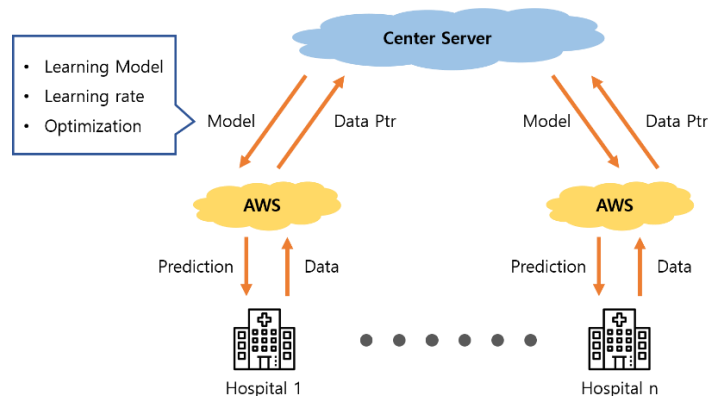


그림 2. 연합학습 아키텍처

2.3.3 연합학습 모델

FL 아키텍처 개발을 위해서는 두 가지의 모델이 구축되어야 한다. 클라이언트들이 로컬 데이터를 사용하여 학습하게 되는 로컬 모델과 클라이언트로부터 수신된 가중치를 취합할 Aggregation 모델이 필요하다.

본 연구에서는 로컬 모델로 MLP 기반 뇌연령 예측모델을 구축하여 연합학습의 성능을 테스트하였다. MLP 모델은 총 3개의 은닉층을 이루고 있으며, 첫번째 은닉층에서는 시그모이드 함수를 사용하였다(그림 3)[10]. 모델 최적화는 L1 Loss function을 바탕으로 Adam(learning rate = 0.003)을 사용하였다. Iteration은 100으로 설정하여 Non-FL 모델과 FL 모델 모두 동일한 조건하에 학습을 진행하였다.

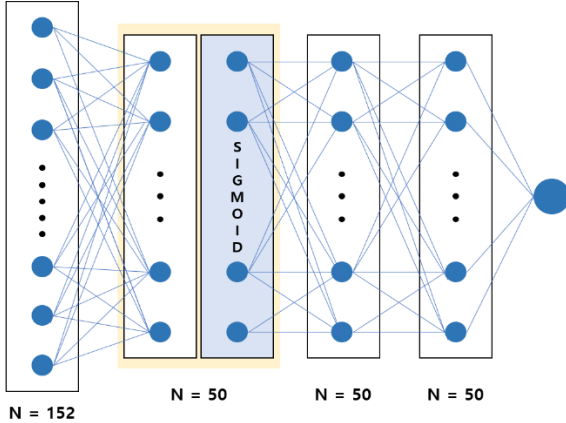


그림 3. 로컬 모델 MLP 아키텍처

각 클라이언트로부터 수신된 가중치들을 취합하기 위해 Aggregation 모델을 구축하였다. 모든 로컬에서 학습된 모델의 가중치값의 평균값을 라운드의 최종 글로벌 가중치로 계산하여 클라이언트에게 배포하는 Federated Averaging 알고리즘을 사용하였다. 중앙 서버는 각 클라이언트에게 45,600 개의 파라미터들을 수신 받아 평균화를 통해 글로벌 가중치를 만들어 배포하고, 총 3 라운드의 Aggregation 을 수행하였다.

3. 연구 결과

3.1 정규분포 데이터 기반 FL 성능비교

정규분포 데이터 상황에서 동일한 조건에서 학습된 FL 모델은 클라이언트 1의 경우 Non-FL 모델 대비 MAE값은 약 2.7% 낮아졌고, 클라이언트 2는 Non-FL 대비 MAE값은 약 1.3% 올라가는 결과를 보였다. R값의 경우, 클라이언트 1과 2에서 모두 유사한 성능을 보였다.

3.2 편향분포 데이터 기반 FL 성능비교

편향분포 데이터 상황에서 동일한 조건에서 학습된 FL 모델은 클라이언트 1의 경우 Non-FL 모델 대비 MAE값은 약 28.7% 낮아졌고, 클라이언트 2는 Non-FL 대비 MAE값 31.7% 낮아져서 성능 개선이 이루어졌다. R값의 경우, 클라이언트 1은 약 5.0% 성능 개선이 이루어졌고, 클라이언트 2는 약 7.4% 성능이 향상되는 것을 확인하였다.

표 2. 분포 및 클라이언트 별 Non-FL 과 FL 성능비교

분포	Client	학습 방법	MAE	R
정규 분포	Client 1	Non-FL	8.02	0.775
		FL	7.80	0.779
	Client 2	Non-FL	7.31	0.863
		FL	7.41	0.859
편향 분포	Client 1	Non-FL	12.42	0.726
		FL	8.58	0.763
	Client 2	Non-FL	11.92	0.775
		FL	8.14	0.833

4. 결론 및 향후 연구

개인정보를 포함하는 데이터들의 활용이 어려워짐에 따라 직접적인 데이터를 공유없이, 모델을 학습시킬 수 있는 FL 패러다임의 중요성이 확대되고 있다. 특히 본 연구에서는 개인정보보호가 중요한 의료데이터를 활용하여 FL을 적용하였다는 점에서 의의가 있다. 본 연구에서는 FL 모델로 학습을 진행할 경우, 독립적인 로컬 모델로 학습할 때 보다 성능이 더 향상되는 것을 확인할 수 있다. 특히 편향된 데이터 분포를 가진 로컬 연합 클라이언트 환경에서 연합학습을 통해 뇌연령 예측의 정확도가 향상되는 것으로 실험적으로 증명했다. 결과적으로, 이 연구는 의료데이터 기반의 FL 성능의 가능성을 확인하는 선행연구로써, 추후 Aggregation 모델 개선, 클라이언트 수 확대를 통한 새로운 FL 연구의 방향성을 제시할 수 있을 것으로 기대한다.

5. 참고문헌

- [1] 이승민, 연합학습 기술 동향 및 산업적 시사점, 기술정책 트렌드(ETRI), 2020
- [2] Keith Bonawitz et al., "Practical secure aggregation for federated learning on user-held data", NIPS Workshop, 2016
- [3] W. Li et al., "Privacy-preserving Federated Brain Tumour Segmentation", 2019
- [4] Jonsson et al., "Brain age prediction using deep learning uncovers associated sequence variants", nature communications, 10:5409, 2019
- [5] <https://www.cam-can.org/>
- [6] <https://surfer.nmr.mgh.harvard.edu/>
- [7] <https://blog.openmined.org/tag/pysyft/>
- [8] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011
- [9] H. Brendan McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data", JMLR: volume 54., 2017
- [10] Hualou Liang et al., "Investigating systematic bias in brain age estimation with application to post-traumatic stress disorders", Hum Brain Mapp, 40: 3143-3152, 2019