

# 연합학습 시 글로벌 모델 성능 향상을 위한 기여도 기반 집계 알고리즘

전석환\*<sup>1</sup>, 이주원\*<sup>2</sup>, 방준일\*<sup>3</sup>, 김화종\*\*

## Contribution-based Aggregation Algorithm to Improve Global Model Performance During Federated Learning

Sukhwan Jeon\*<sup>1</sup>, Joowon Lee\*<sup>2</sup>, Junil Bang\*<sup>3</sup>, and Hwajong Kim\*\*

이 논문은 2023년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과임(2022RIS-005)

### 요약

연합학습은 데이터 공유 문제의 해결책으로 떠오르고 있으나, 클라이언트의 데이터 분포가 불균형한 환경에서 특정 클라이언트가 글로벌 모델의 성능을 하락시키는 클라이언트 데이터 편향 문제가 존재한다. 본 논문에서는 이러한 글로벌 모델 학습 편향을 해결하기 위해 각 클라이언트의 기여도를 측정하고, 기여도에 비례하게 모델 파라미터를 가중 평균하여 기여도가 낮은 클라이언트(편향된 데이터를 가지고 있는 클라이언트)가 글로벌 모델에 주는 영향을 줄이는 방법을 제안한다. 실험 결과 FedAVG 대비 IID 상황에서는 동일한 성능의 글로벌 모델을 집계하였고, NON-IID 상황에서는 정확도가 1.5% 향상된 글로벌 모델을 집계하며 FedAVG 기본 성능은 유지하고 단점만을 보완한 알고리즘임을 증명하였다.

### Abstract

Federated learning is emerging as a solution to the data sharing problem, but in an environment where the data distribution of clients is unbalanced, there is a problem of client data bias in which certain clients degrade the performance of the global model. In this paper, in order to solve this global model learning bias, we propose a method to reduce the impact of low-contribution clients on the global model by measuring the contribution of each client and weighting the model parameters in proportion to the contribution. As a result of the experiment, a global model with the same performance was aggregated in the IID situation compared to FedAVG, and a global model with 1.5% improved accuracy was aggregated in the NON-IID situation, proving that it is an algorithm that maintains the basic performance of FedAVG and complements only the shortcomings.

### Keywords

federated learning, shapley value, contribution, aggregation algorithm, non-iid

\* 강원대학교 컴퓨터정보통신공학과

- ORCID<sup>1</sup>: <https://orcid.org/0000-0002-0882-6347>

- ORCID<sup>2</sup>: <https://orcid.org/0000-0001-8638-4081>

- ORCID<sup>3</sup>: <https://orcid.org/0000-0003-0582-1572>

\*\* 강원대학교 컴퓨터정보통신공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0002-3822-390X>

· Received: Dec. 03, 2023, Revised: Dec. 14, 2023, Accepted: Dec. 17, 2023

· Corresponding Author: Hwajong Kim

Dept. of Computer and Communications Engineering, Kangwon National Univ., Gangwondaehak-gil 1, Chuncheon-si, Gangwon-do, Korea

Tel.: +82-33-250-6323, Email: [hjkim3@gmail.com](mailto:hjkim3@gmail.com)

## I. 서 론

머신러닝 모델의 성능은 데이터의 품질과 양에 비례한다. 이러한 이유로 머신러닝 모델의 성능 향상을 위해 다양하고 많은 데이터를 구축하며 많은 시간과 비용을 소모하고 있지만, 개인정보 및 보안 문제 등으로 데이터를 한곳에 모으는 일은 한계가 존재한다. 이를 개선하기 위해 비실명화된 데이터의 공유를 가능하게 하는 데이터 3법이 통과되었지만, 아직 실용화를 위해서는 해결해야 하는 문제가 여럿 존재한다. 예를 들어, 경쟁 기관 간 협업을 통해 머신러닝 모델 학습데이터를 구축한다고 가정하면, 데이터를 한곳에 모으기 위해 경쟁기관에 데이터를 공유하여 공개해야만 한다. 이러한 이유로 데이터 공유를 통한 학습데이터 구축은 많은 제한조건이 있다.

이에 연합학습(Federated learning)이 해결 방안으로 등장하였다. 연합학습은 개인 데이터를 외부로 전송하지 않고 머신러닝 모델 파라미터만을 공유하는 방식이다. 이를 통해 개인정보가 외부로 유출되지 않고 모든 데이터를 학습한 성능의 글로벌 모델을 모든 클라이언트가 사용할 수 있다. 이러한 장점으로 많은 분야에서 연합학습을 적용하는 연구가 활발히 진행되고 있다[1]-[4].

머신러닝 학습데이터 공유 문제점을 해결한 연합학습에도 한계점이 존재한다. 연합학습에 참여하는 각 클라이언트의 데이터는 데이터의 품질과 양, 타겟(클래스)의 분포가 다를 수 있으며, 이러한 데이터 분포의 불균형은 글로벌 모델의 학습이 편향되는 문제를 야기한다. 이러한 편향을 해결하기 위한 글로벌 모델 집계 방식인 FedAVG가 현재 가장 많이 사용되고 있지만, FedAVG는 데이터양에 비례하여 가중 평균을 하는 방식이기에, 데이터의 타겟 및 품질에 대한 분포는 고려하지 않는다.

본 논문에서는 이러한 글로벌 모델 학습 편향을 해결하기 위해 각 클라이언트의 기여도를 측정하고, 기여도에 비례하게 모델 파라미터를 가중 평균하여 기여도가 낮은 클라이언트(편향된 데이터를 가지고 있는 클라이언트)가 글로벌 모델에 주는 영향을 줄이는 방법을 제안한다.

## II. 관련 연구

### 2.1 연합학습

연합학습은 2016년 구글에서 키보드 글자 입력 예측(Gboard)을 위해 도입한 학습 방법[5]으로, 데이터를 직접 공유하지 않고, 각 클라이언트가 로컬 디바이스에서 학습한 모델의 파라미터만을 공유하고 집계하여 마치 모든 클라이언트의 데이터를 모두 학습한 머신러닝 모델의 성능과 유사한 글로벌 모델을 만드는 학습 기법이다. 집계된 글로벌 모델은 다시 모든 클라이언트로 배포되며, 모든 참여 클라이언트는 높은 성능의 글로벌 모델을 사용할 수 있다. 현재까지 연합학습을 통해 집계된 글로벌 모델의 성능이 전체 데이터를 한곳에 모아 학습한 모델과 유사한 성능을 낼 수 있음을 입증한 논문이 다수 발표되었다[6]-[12].

연합학습은 개별 데이터를 가지고 있는 다수의 클라이언트와, 수집된 클라이언트들의 모델 파라미터를 집계하여 글로벌 모델을 만드는 서버로 나뉘어 동작한다. 클라이언트의 로컬 모델을 학습하고 글로벌 모델을 집계하는 과정을 여러 번 반복하여 글로벌 모델의 성능을 점차 향상시키며, 이렇게 반복되는 과정을 하나의 round라고 한다. 첫 round는 각 클라이언트에 배포되는 모델을 초기화하는 작업이 포함되며, 이후의 round는 클라이언트 로컬 학습, 집계 및 배포만 반복된다. 자세한 동작 과정은 아래 설명과 같으며 그림 1로 표현하였다.

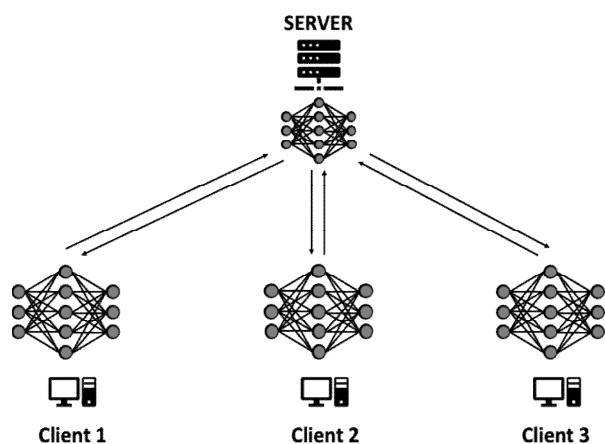


그림 1. 연합학습 개념도

Fig. 1. Federated learning concept map

#### - 첫 round 동작

1. 공통 모델 초기화 : 각 클라이언트는 동일한 초기 모델을 학습한 후 집계해야 하기에 서버 측에서 초기 모델을 생성하여 초기 글로벌 모델로 설정한다.

2. 최소 참여 클라이언트 연결 : 연합학습 진행을 위해 사전에 설정하는 수치 중 최소 클라이언트 수가 있다. 설정한 최소 클라이언트 수만큼 통신이 연결되기 전까지 대기하고, 연결이 완료되면 연합학습을 실행한다.

#### - 모든 round 동작

1. 모델 배포 : 클라이언트 로컬 학습을 진행하기 위해 모든 클라이언트는 서버로부터 동일한 글로벌 모델을 수신받는다.

2. 클라이언트 로컬 학습 : 서버로부터 수신받은 글로벌 모델을 클라이언트의 개별 데이터로 학습을 진행한다.

3. 클라이언트 모델 파라미터 공유 : 클라이언트의 로컬 학습이 완료되면 서버로 학습이 완료된 모델의 파라미터를 전송한다.

4. 글로벌 모델 집계 : 클라이언트로부터 학습이 완료된 모델을 파라미터를 모두 수신받은 후 모델 파라미터를 선택한 집계 알고리즘에 맞게 집계하여 글로벌 모델을 만든다.

## 2.2 FedAVG

FedAVG(Federated Averaging)는 연합학습에서 사용되는 주요 집계 알고리즘[5] 중 하나로 현재 연합학습에 가장 많이 사용되는 집계 알고리즘이다. 클라이언트로부터 받은 모델 파라미터를 각 클라이언트가 소유한 데이터의 개수에 비례하여 가중 평균하는 방식이다. FedAVG의 수식은 아래 식 (1)과 같으며,  $K$ 는 클라이언트 번호,  $k$ 는 전체 클라이언트의 개수,  $w$ 는 모델 가중치,  $w^k$ 는  $k$ 번 클라이언트 로컬 모델 가중치,  $n$ 은 모든 클라이언트 데이터 개수의 합,  $n_k$ 는  $k$ 번 클라이언트의 데이터 개수를 의미한다.

$$w \leftarrow \sum_{k=1}^K \frac{n_k}{n} w^k \quad (1)$$

## 2.3 Shapley value

연합학습 참여 클라이언트의 기여도를 계산하는 방법으로 Shapley value를 사용하는 연구가 많이 진행되고 있다. 2012년 노벨경제학상을 수상한 Lloyd Shapley의 이론[13]인 Shapley value는 협력형 게임이론에서 보수지급(Payoff) 방법으로 제안되었는데, 상호 협력하기로 한 게임 참여자가 전체적인 성능 향상에 기여한 정도를 점수화하는 방법이다. Shapley value를 계산하는 과정을 간단하게 설명하면 아래와 같다.

1. 참여 클라이언트가 생성할 수 있는 모든 부분집합을 만들고, 모든 부분집합에 대해 부분집합에 포함된 클라이언트만으로 연합학습을 진행하여 최종 모델의 성능을 추출한다.

2. 부분집합 중,  $n$ 번 클라이언트가 포함된 부분집합과 포함되지 않은 부분집합의 성능 차이를 통해  $n$ 번 클라이언트가 참여함으로써 얻을 수 있는 성능 차이를 계산한다.

3. 모든 클라이언트에 대해 2번 과정을 반복하여 각 클라이언트의 Shapley Value를 얻는다.

위 과정의 Shapely Value  $\rho(i)$ 의 수식은 아래 수식 (2)와 같으며,  $N$ 은 참여 클라이언트의 수,  $S$ 는 참여 클라이언트의 부분집합,  $i$ 는 특정 클라이언트,  $v(S)$ 는 부분집합  $S$ 의 성능이다. 본 논문의 실험에서는 부분집합  $S$ 의 성능으로 accuracy를 적용하였다.

$$\rho(i) = \sum_{S \subset M \setminus \{i\}} \frac{|S|!(N-|S|-1)!}{N!} \{v(S \cup i) - v(S)\} \quad (2)$$

## III. 기여도 기반 집계 알고리즘

연합학습 시 클라이언트의 데이터가 글로벌 모델에 부정적인 영향을 끼치는 이유는 데이터의 개수 외로 데이터의 품질, 클래스 불균형, 잘못된 레이블링 등 다양한 이유가 존재한다. 2.2에서 설명한 FedAVG의 경우 각 클라이언트의 데이터 개수에 대한 불균형은 고려하였지만, 데이터 품질, 클래스 불균형 등의 다른 사항들은 고려하지 않는다. 따라서 FedAVG를 사용하는 연합학습의 경우 데이터의 분포 및 품질이 좋지 않은 클라이언트가 참여하여 글로벌 모델의 성능을 낮추는 상황도 발생한다.

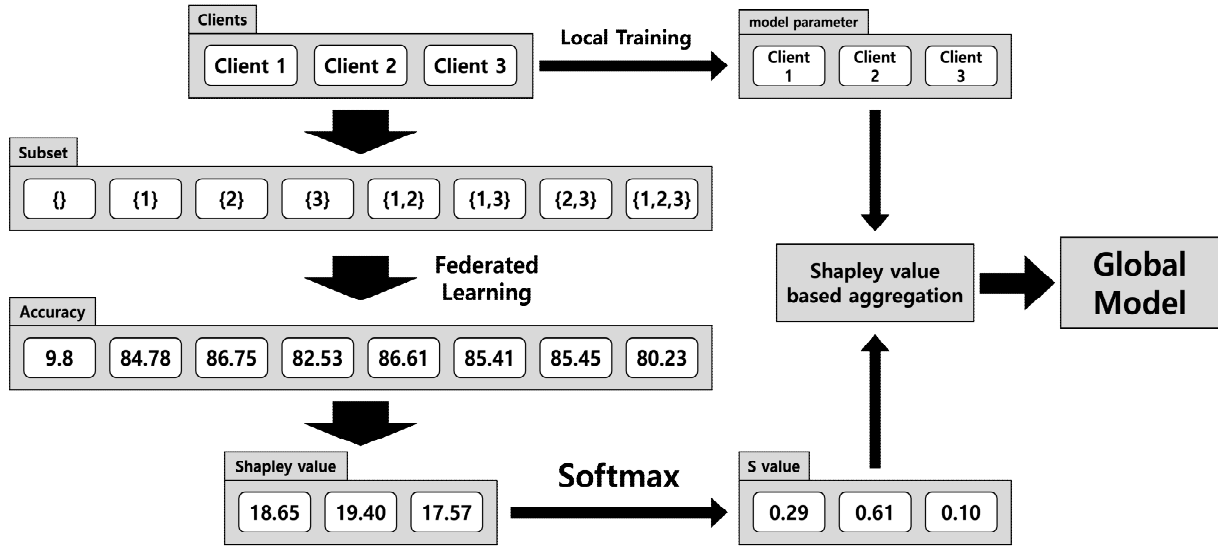


그림 2. round 별 s값 계산 및 기여도 기반 집계 과정  
 Fig. 2. Calculation of s value for each round and contribution-based aggregation process

본 논문에서는 위와 같은 문제점을 해결하기 위해 글로벌 모델 집계 시 Shapley value 기반의 가중치를 추가하여 기여도가 낮은 클라이언트가 글로벌 모델에 주는 영향을 줄이도록 하였다. Shapley value가 추가된 모델 집계 수식은 아래 식 (3)과 같으며,  $K$ 는 클라이언트 번호,  $k$ 는 전체 클라이언트의 개수,  $w$ 는 모델 가중치,  $w^k$ 는  $k$ 번 클라이언트 로컬 모델 가중치,  $s_k$ 는  $k$ 번 클라이언트의 기여도를 의미한다. 기여도  $s$ 는 0 ~ 1 사이의 값을 가지고, 모든 클라이언트 기여도 합은 1이 되도록 Shapley value에 softmax 함수를 사용하였다.

$$w \leftarrow \sum_{k=1}^K \frac{s_k}{1} w^k \quad (3)$$

2.3에서 설명한 연합학습의 동작 과정은 모든 round마다 글로벌 모델을 집계한다. 즉, 식 (3)을 적용하여 글로벌 모델을 집계하기 위해선 모든 round에서 Shapley value를 계산해야 하며, softmax를 통해  $s$ 값을 구해야 한다. 한 round에서  $s$ 값을 구하고 글로벌 모델을 집계하는 과정은 아래와 같으며 그림 2로 표현하였다.

1. round가 시작되면 참여 클라이언트의 모든 부분 집합을 생성한다.

2. 각 클라이언트로 글로벌 모델을 배포하는 단계에서 생성된 부분집합 중 하나의 집합을 모든 클라이언트로 전송한다.
3. 각 클라이언트는 수신받은 부분집합에 자신이 포함되어 있는지 확인한 후, 포함되어 있다면 정상적으로 연합학습을 진행하고, 포함되어 있지 않다면 학습을 건너뛰고 데이터 개수를 0으로 변경하여 서버로 전송한다.
4. 서버는 모든 클라이언트로부터 모델 파라미터를 수신받으면 FedAVG 알고리즘을 적용하여 임시 글로벌 모델을 집계한다. 이때 FedAVG 알고리즘 수식에 따라 데이터 개수가 0인 클라이언트는 집계에 영향을 미치지 않는다. 즉, 부분집합에 포함되지 않은 클라이언트의 모델 파라미터는 집계되지 않는다.
5. 집계된 글로벌 모델의 성능을 추출한다.
6. 2~5의 과정을 모든 부분집합에 적용되도록 반복하여 모든 부분집합에 대한 임시 글로벌 모델 성능을 추출한다.
7. 모든 부분집합에 대한 성능을 통해 각 클라이언트의 Shapleyvalue를 계산하고 softmax 함수를 통해  $s$ 값을 계산한다.
8. 각 클라이언트의  $s$ 값을 통해 식 (3)을 적용하여 해당 round의 최종 글로벌 모델을 집계한다.

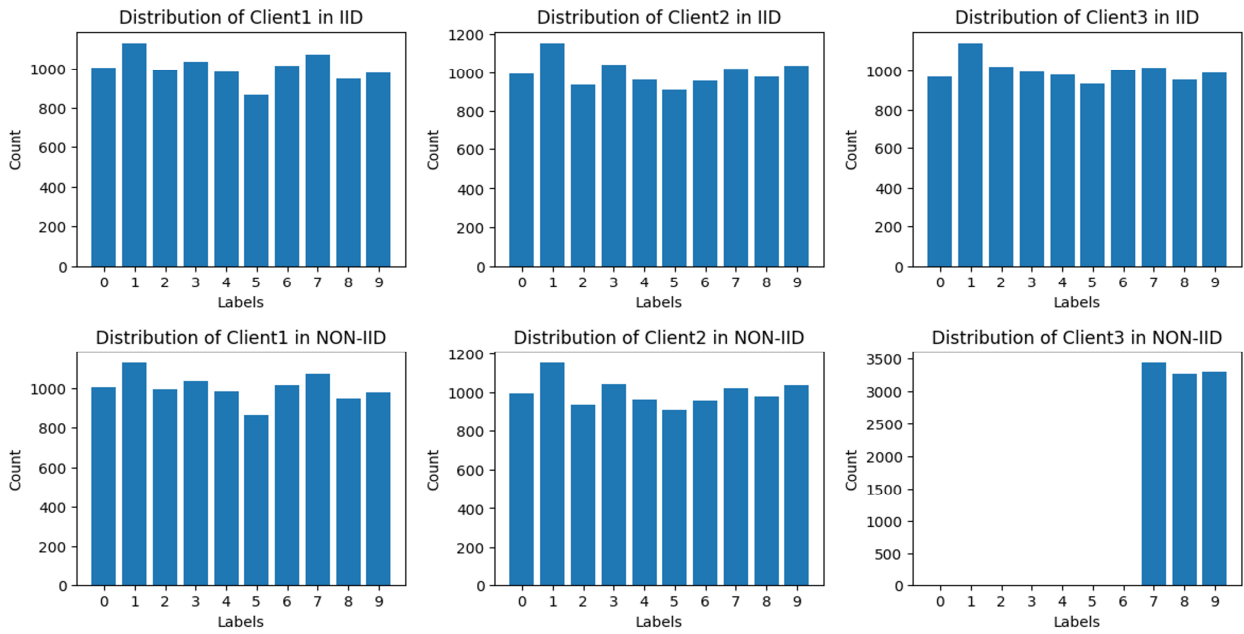


그림 3. 상황 및 클라이언트 별 데이터 분포  
Fig. 3. Data distribution by case and client

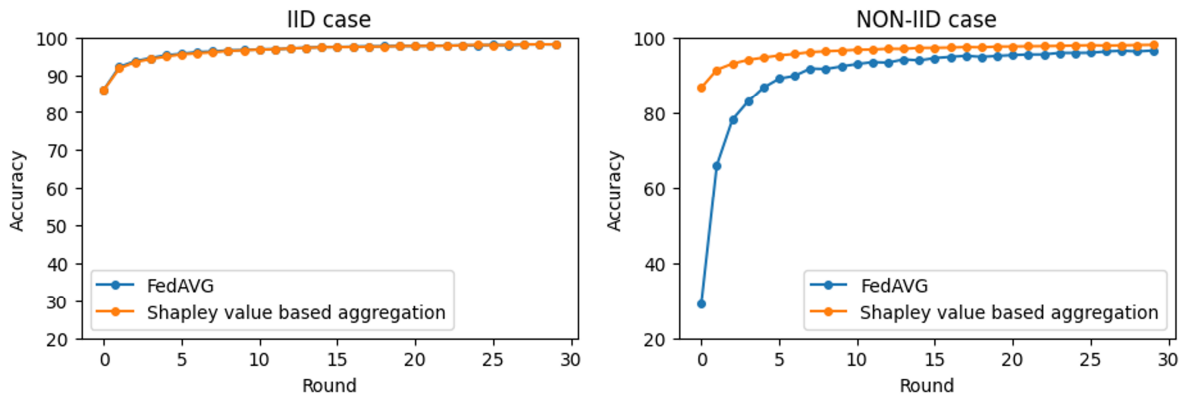


그림 4. IID 상황과 NON-IID 상황의 집계 결과 Accuracy 비교  
Fig. 4. Accuracy comparison of aggregate results between IID and NON-IID situations

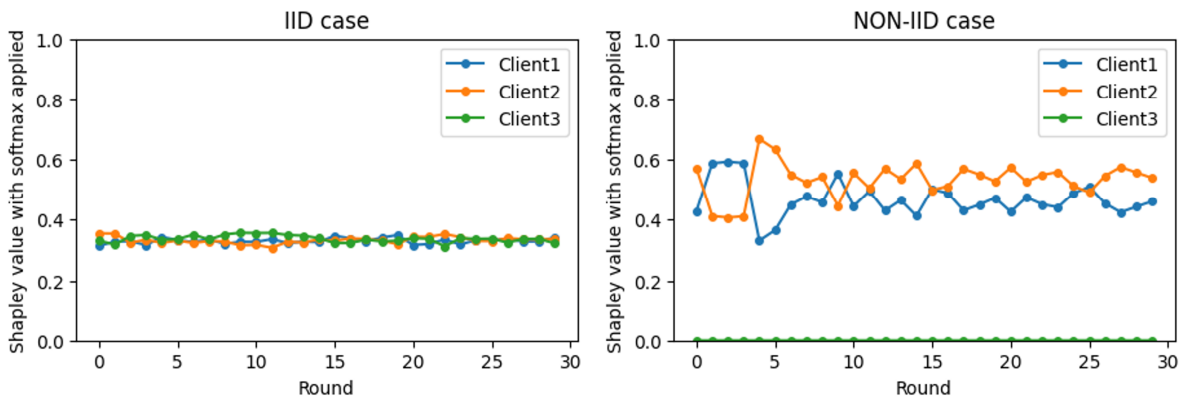


그림 5. round별 IID 상황과 NON-IID 상황의 s값 비교  
Fig. 5. Comparison of s values between IID situation and NON-IID situation by round

위와 같이 기여도에 기반하여 글로벌 모델을 집계하면, 글로벌 모델의 성능에 좋은 영향을 끼치는 클라이언트의 영향이 커진다. 이로 인해 글로벌 모델 성능에 영향을 주는 데이터 개수, 품질, 클래스 불균형 외로 이전에 고려하지 못한 성능에 대한 모든 사항이 집계에 반영된다. 또한, 연합학습 초기에는 글로벌 모델에 큰 기여를 하지 못하였지만, 모델이 학습되고 손실값이 수렴되며 뒤늦게 글로벌 모델에 좋은 영향을 주는 클라이언트도 존재한다. 이러한 클라이언트 또한 모든 round마다 기여도를 측정하고 집계하기에, round가 충분히 진행되면 글로벌 모델에 영향을 줄 수 있다.

## IV. 실험 및 결과

### 4.1 실험

실험에 사용한 데이터는 MNIST 데이터셋 [14]-[16]이다. MNIST 데이터셋은 0부터 9까지의 손글씨 이미지로 이루어진 벤치마크 데이터셋으로 10개의 클래스가 각각 7,000개씩 총 70,000개의 데이터로 구성된다. 일반적으로 70,000개의 데이터를 60,000개의 학습데이터와 10,000개의 검증데이터로 나누어서 실험하지만, 실험 결과를 쉽게 비교할 수 있도록 각 클라이언트당 10,000개의 학습데이터와 3,333개의 검증 데이터로 분할 하여 총 3개의 클라이언트로 실험을 진행하였다. 또한, 실험 환경을 IID(Independent and Identically Distributed)와 NON-IID 두가지 환경으로 구성하였다. 두 상황 모두 각 클라이언트의 데이터 개수는 동일하며 분포는 그림 3과 같다. 실험에 사용한 모델은 Pytorch CNN 모델을 사용했으며, 연합학습 구현은 Adap사에서 공개한 Django 및 gRPC 통신 기반 연합학습 프레임워크인 Flower를 사용하였다[17].

### 4.2 실험 결과

그림 4에는 FedAVG와 기여도 기반 집계의 글로벌 모델 성능을 IID 상황과 NON-IID 상황을 나누어 비교한 그래프이다. 모든 클라이언트의 데이터 분포가 균일한 IID 상황에서는 FedAVG와 기여도 기반 집계의 차이가 크게 나지 않았으며, 20 round 이후

부터는 기여도 기반 집계 방식이 미세하게 성능이 높은 경우도 존재했다. IID 상황에서의 최고 정확도는 FedAVG와 기여도 기반 집계 방식 모두 98.18%로 동일하였다.

반대로 3번 클라이언트 데이터의 label이 7,8,9만 존재하는 클래스 불균형 상황에서는 첫 round부터 FedAVG보다 높은 성능을 보였다. 이후 30 round까지 FedAVG보다 높은 성능을 보이며 기여도 기반 집계 방식이 의미 있게 동작함을 증명하였다. NON-IID 상황에서 FedAVG와 기여도 기반 집계 방식의 최고 정확도는 각각 96.62%, 98.13%로 기여도 기반 집계 방식의 정확도가 약 1.5% 더 높았다. IID 상황 및 NON-IID 상황의 집계 방식별 모델 최고정확도를 아래 표 1로 정리하였다.

표 1. 집계 방식 별 모델 최고 정확도

Table 1. Highest model accuracy by aggregation algorithm

	Aggregation algorithm	Accuracy
IID	FedAVG	98.18 %
	Contribution-based aggregation	98.18 %
NON-IID	FedAVG	96.62 %
	Contribution-based aggregation	98.13 %

그림 5에는 round 별 각 클라이언트의  $s$  값 변화를 표시한 그래프이다. IID 상황의 경우 모든 클라이언트가 비슷한 기여도를 유지하는 것을 볼 수 있다. NON-IID 상황에서는 3번 클라이언트의 기여도가 0에 가깝게 나오면서 글로벌 모델에 영향을 거의 주지 않았음을 확인할 수 있다. 이는 3번 클라이언트는 오히려 글로벌 모델에게 악영향을 주는 것으로 해석할 수 있다.

특이한 점은 round가 진행될수록 모든 클라이언트의 기여도가 비슷해 지는 쪽으로 변화하는 것을 볼 수 있다. 이는 글로벌 모델의 성능이 충분히 높아짐에 따라 각 클라이언트의 로컬 모델이 글로벌 모델에 주는 영향이 작아지기 때문으로 추측한다.

## V. 결론 및 향후 과제

연합학습은 데이터 공유 문제의 해결책으로 떠오

르고 있으나 클라이언트의 데이터 불균형 문제를 해결해야만 실용화에 가능하다. 본 논문은 이러한 클라이언트 데이터 불균형에 따른 글로벌 모델 성능 저하를 막기 위해 기존 FedAVG 대신 기여도에 가중 평균하여 클라이언트 로컬 모델의 파라미터를 집계하는 기여도 기반 집계 알고리즘을 제안하였다. 검증을 위해 각 클라이언트의 데이터 분포를 IID 상황과 NON-IID 상황으로 구분하여 실험하였다. 그 결과 IID 상황에서는 기존 FedAVG와 동일한 성능의 글로벌 모델을 집계하였고, NON-IID 상황에서는 FedAVG 대비 정확도가 1.5% 향상된 글로벌 모델을 집계하며 FedAVG 기본 성능은 유지하고 단점만을 보완한 알고리즘임을 증명하였다.

다만, 기여도 기반 집계에 사용한 Shapley Value는 계산 복잡도가 높은 문제점이 존재한다.  $n$ 개 클라이언트의 부분 집합 개수는  $2^n$ 개로 클라이언트의 수가 증가함에 따라 학습 시간이 지수적으로 늘어나게 된다. 이를 개선하기 위해 몬테카를로 기법과 같은 Shapley value의 계산량을 줄이는 연구가 필요하다.

본 논문의 실험에서는 일부 클라이언트 데이터의 label 분포를 다르게 하여 NON-IID 환경을 구성하였지만, label 오류가 있는 데이터 혹은 노이즈가 있는 데이터 등의 품질이 낮은 데이터를 보유한 클라이언트와의 연합학습 실험도 검증이 필요하다. 따라서 Shapley value의 계산량 문제 및 다양한 실험 환경을 고려한 연구를 본 논문의 향후 과제로 한다.

## References

- [1] W. Li, et al., "Privacy-preserving federated brain tumour segmentation", MLMI 2019: Machine Learning in Medical Imaging, Vol. 11861, pp. 133-141, Oct. 2019. [https://doi.org/10.1007/978-3-030-32692-0\\_16](https://doi.org/10.1007/978-3-030-32692-0_16).
- [2] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation", BrainLes 2018: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Vol. 11383, pp. 92-104, Jan. 2019. [https://doi.org/10.1007/978-3-030-11723-8\\_9](https://doi.org/10.1007/978-3-030-11723-8_9).
- [3] M. J. Sheller, et al., "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data", Nature Scientific Reports, Vol. 10, No. 12598, Jul. 2020. <https://doi.org/10.1038/s41598-020-69250-1>.
- [4] J. Bang, S. Hong, S. Jeon, J. Lee, and H. Kim, "Federated Learning Study using Motion Recognition Model based on Joint Data", Journal of KIIT, Vol. 21, No. 3, pp. 39-48, Mar. 2023. <http://dx.doi.org/10.14801/jkiit.2023.21.3.39>.
- [5] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", Artificial intelligence and statistics, PMLR, Vol. 54, pp. 1273-1282, Apr. 2017.
- [6] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, and S. Bakas, "Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation", BrainLes 2018: Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Vol. 11383, pp. 92-104, Jan. 2019. [https://doi.org/10.1007/978-3-030-11723-8\\_9](https://doi.org/10.1007/978-3-030-11723-8_9).
- [7] W. Li, et al., "Privacy-preserving federated brain tumour segmentation", MLMI 2019: Machine Learning in Medical Imaging, Vol. 11861, pp. 133-141, Oct. 2019. [https://doi.org/10.1007/978-3-030-32692-0\\_16](https://doi.org/10.1007/978-3-030-32692-0_16).
- [8] H. K. Ko, "A Study Personal Information Protection Technique for XML-based Electronic Medical Record", Journal of KIIT, Vol. 12, No. 5, pp. 185-191, May 2014. <https://doi.org/10.14801/kiitr.2014.12.5.185>.
- [9] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition", In 2020 International Conference on Artificial Intelligence in Information and Communication(ICAIC), Fukuoka, Japan, pp. 362-366, Feb. 2020. <https://doi.org/10.1109/ICAIC48513.2020.9065078>.

[10] M. J. Sheller, et al., "Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data", Nature Scientific Reports, Vol. 10, No. 12598, Jul. 2020. <https://doi.org/10.1038/s41598-020-69250-1>.

[11] H. J. Kim, S. E. Hong, and K. J. Cha, "seq2vec: analyzing sequential data using multi-rank embedding vectors", Electronic Commerce Research and Applications, Vol. 43, Sep. 2020. <https://doi.org/10.1016/j.elerap.2020.101003>.

[12] H. Cho and S. M. Yoon, "Divide and conquer-based 1D CNN human activity recognition using test data sharpening", Sensors, Vol. 18, No. 4, Apr. 2018. <https://doi.org/10.3390/s18041055>.

[13] L. S. Shapley, "The theory of stable allocations and the practice of market design. The Nobel Prize in Economics 2012 for Alvin E. Roth", Contributions to Science, Vol. 11, pp. 103-112, Jan. 2015. <https://doi.org/10.2436/20.7010.01.218>.

[14] L. Bottou, et al., "Comparison of classifier methods: a case study in handwritten digit recognition", Proceedings of the 12th IAPR International Conference on Pattern Recognition, Jerusalem, Israel, Vol. 3, pp. 77-82, Oct. 1994. <https://doi.org/10.1109/ICPR.1994.576879>.

[15] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition", Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278-2324, Nov. 1998. <https://doi.org/10.1109/5.726791>.

[16] Y. LeCun, C. Cortes, and C. Burges, MNIST handwritten digit database, 1998. <http://yann.lecun.com/exdb/mnist/index.html> [accessed : Nov. 29, 2023]

[17] D. J. Beutel, et al., "Flower: A Friendly Federated Learning Research Framework", arXiv preprint arXiv:2007.14390, Jul. 2020. <https://doi.org/10.48550/arXiv.2007.14390>.

저자소개

전 석 환 (Sukhwan Jeon)



2022년 : 강원대학교  
기계융합공학과(공학사)  
2022년 ~ 현재 : 강원대학교  
컴퓨터정보통신공학과 석사과정  
관심분야 : 기계학습, 딥러닝,  
빅데이터

이 주 원 (Joowon Lee)



2017년 ~ 2023년 : 강원대학교  
정보통신공학전공(공학사)  
2023년 ~ 현재 : 강원대학교  
컴퓨터정보통신공학과 석사과정  
관심분야 : 빅데이터, 기계학습,  
인공지능

방 준 일 (Junil Bang)



2018년 9월 ~ 2020년 8월 :  
강원대학교  
컴퓨터정보통신공학과(공학석사)  
2020년 9월 ~ 현재 : 강원대학교  
컴퓨터정보통신공학과 박사과정  
관심분야 : 데이터분석, 머신러닝

김 화 중 (Hwajong Kim)



1984년 3월 : KAIST  
전기및전자공학과(공학석사)  
1988년 3월 : KAIST  
전기및전자공학과(공학박사)  
1988년 3월 ~ 현재 : 강원대학교  
컴퓨터정보통신공학과 정교수  
관심분야 : 데이터 통신,  
컴퓨터네트워크, 네트워크 프로그래밍, 빅데이터