Smart Media Journal / Vol.14, No.6 / ISSN:2287-1322
https://dx.doi.org/10.30693/SMJ.2025.14.6.42

비전 기반 멀티태스크 학습을 위한 연합학습 알고리즘의 성능 분석

(Performance Analysis of Federated Learning Algorithms for Vision-Based Multi-Task Learning)

강태욱*, 박철영, 이명배** 신창선*****

(Tae Wook Kang, ChulYoung Park, MyengBae Lee, Chang Sun Shin)

요 약

본 논문은 비전 기반 멀티태스크 학습으로의 확장을 염두에 두고, 연합학습(Federated Learning) 환경에서 대표적인 알고리즘인 FedSGD(Federated Stochastic Gradient Descent), FedAVG(Federated Averaging), FedRep(Federated Representation Learning)의 성능을 비교·분석한다. 데이터 프라이버시와 통신 효율성을 고려한 분산 학습 구조인 연합학습은 실제 환경에서 클라이언트 간 데이터 불균형(Non-IID) 문제가 빈번하게 발생한다. 본 논문에서는 FEMNIST(Federated Extended MNIST) 데이터셋을 Non-IID 방식으로 분할하여 각 알고리즘의 정확도, 통신 효율성, 손실을 정량적으로 평가하였다. 실험 결과, FedAVG는 전반적으로 우수한 성능과 안정성을 보였으며, FedRep은 사용자 맞춤형 학습에 강점을 보였다. FedSGD는 통신량은 적었으나 정확도 면에서는 상대적으로 취약하였다. 이러한 결과는 향후 분류뿐만 아니라 객체 탐지, 의미론적 분할 등 멀티태스크 학습 환경으로 연합학습을 확장할 때 알고리즘 선택의 기초 자료로 활용될 수 있다.

■ 중심어 : 연합학습 ; Non-IID ; FedAVG ; FedSGD ; FedRep

Abstract

This study aims to evaluate the performance of key Federated Learning (FL) algorithms—FedSGD(Federated Stochastic Gradient Descent), FedAVG(Federated Averaging), and FedRep(Federated Representation Learning)—with a view toward future expansion into vision-based multitask learning. Federated learning, a distributed training paradigm that prioritizes data privacy and communication efficiency, faces significant performance degradation in real-world settings due to the frequent presence of Non-IID (non-identically distributed) data across clients. In this work, we partitioned the FEMNIST(Federated Extended MNIST) dataset in a Non-IID manner and quantitatively assessed the convergence speed, accuracy, and communication efficiency of each algorithm. Experimental results show that FedAVG consistently delivers strong performance and stability, while FedRep demonstrates advantages in personalized learning. FedSGD, despite requiring fewer communications, exhibited notable limitations in accuracy. These findings provide valuable insights for selecting appropriate algorithms when expanding FL to multitask scenarios, such as object detection and semantic segmentation, beyond simple classification tasks.

■ keywords : Federated Learning ; Non IID ; FedAVG ; FedSGD ; FedRep

Ⅰ. 서 론

최근 인공지능과 기계학습 기술이 급속도로 발전함에 따라, 자율주행, 헬스케어, 금융, IoT 등 다양한 분야에서 인공지능 기반 서비스가 상용화되고 있다. 이와 함께, 대규모 데이터를 활용한

* 준회원, 순천대학교 정보통신공학전공
** 정회원, 순천대학교 인공지능공학과
*** 종신회원, 순천대학교 인공지능공학과
이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2024-00407739).
접수일자 : 2025년 04월 21일 게재확정일 : 2025년 05월 25일
교신저자 : 신창선 e-mail : csshin@scnu.ac.kr

중앙 집중식 학습 방식은 높은 성능을 보장하지만, 개인정보 유출 위험과 과도한 연산 및 통신 비용이라는 한계에 직면하고 있다[1]. 이러한 문제는 특히 데이터 주권과 프라이버시 보호가 강조되는 사회적 흐름 속에서 더욱 부각되고 있으며, 유럽의 GDPR(General Data Protection Regulation), 미국의 CCPA(California Consumer Privacy Act) 등 글로벌 개인정보 보호 법제가 이를 촉진하고 있다[2].

이러한 환경적 제약을 해결하기 위해 제안된 연합학습(Federated Learning, FL)은 데이터를 로컬 장치에 보관한 채, 학습된 모델 파라미터만 중앙 서버와 공유하는 방식으로, 데이터 유출 없이 협력적 학습이 가능한 분산 학습 기술이다. Google의 Gboard, 의료 기관 간 진단 모델 공동 학습 등 실제 사례를 통해 FL은 실용성과 확장성을 동시에 입증하고 있다.

그러나 연합학습은 클라이언트 간 데이터 분포가 상이한 Non-IID 환경에서 성능 저하와 학습 불안정성이 발생하는 구조적 한계를 가진다. 이를 극복하기 위한 다양한 알고리즘이 제안되었으며, 본 논문은 이 중 대표적인 알고리즘인 FedSGD(Federated Stochastic Gradient Descent), FedAVG(Federated Averaging), FedRep(Federated Representation Learning)에 주목하여 모델 손실, 통신 효율성, 모델 정확도의 세 가지 지표를 기준으로 성능을 비교·분석한다.

실험에는 다양한 태스크에 확장 가능한 FEMNIST (Federated Extended MNIST) 데이터셋을 사용하였으며, Non-IID 환경을 가정하여 알고리즘별 학습 특성과 제한점을 정량적으로 평가한다. 본 논문은 단일 분류 작업 기반의 실험 결과를 바탕으로, 향후 객체 탐지(Object Detection), 의미론적 분할(Semantic Segmentation) 등 비전 기반 멀티태스크 학습 환경으로의 확장을 염두에 둔 초기 기반 연구로서, 멀티태스크 연합학습 구조 설계 시 고려해야 할 성능 기준을 도출하고자 한다.

Ⅱ. 관련 연구

가. 연합학습 개념

연합학습은 다수의 클라이언트가 각자의 로컬 데이터로 모델을 학습하고, 그 결과를 서버에 전달하여 글로벌 모델을 공동으로 학습하는 분산형 학습 방식이다. 각 클라이언트는 데이터를 외부에 노출하지 않으며, 중앙 서버는 모델 파라미터(주로 가중치 또는 gradient)를 수집·평균화하는 역할을 수행한다. 이는 데이터 프라이버시를 보호하면서도 대규모 분산 환경에서 협력적 AI 학습을 가능하게 하여, 실제로 헬스케어, 스마트 디바이스, 금융 서비스 등에서 활용되고 있다.

연합학습의 일반적인 프로세스는 다음과 같다:

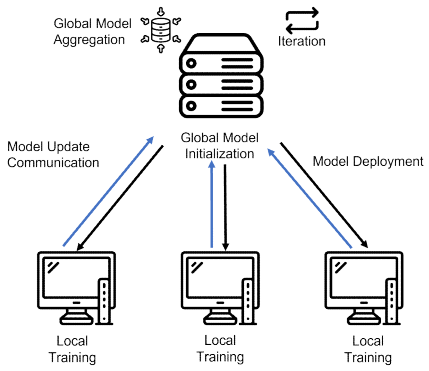

그림 1. 연합학습 프로세스

Step.

① 글로벌 모델 초기화 (Global Model Initialization): 중앙 서버에서 기본 모델을 생성

② 모델 배포 (Model Deployment): 클라이언트(로컬 장치)로 모델을 배포.

③ 로컬 모델 학습 (Local Training): 각 클라이언트는 자신의 데이터를 이용하여 모델 학습.

④ 모델 업데이트 전송 (Model Update Communication): 각 클라이언트에서 로컬 모델의 학습 결과를 서버로 전송.

⑤ 글로벌 모델 집계 (Global Model Aggregation): 중앙 서버에서 여러 클라이언트의 모델을 통합하여 새로운 글로벌 모델 생성.

⑥ 반복 학습 (Iteration): 글로벌 모델을 클라이언트로 다시 배포하여 학습을 지속.

나. 연합학습의 유형

연합학습은 데이터의 특성(feature) 및 사용자(user) 구성 방식에 따라 크게 수평(Horizontal), 수직(Vertical), 연합 전이(Federated Transfer) 학습으로 구분된다.

표 1은 이러한 세 가지 연합학습 유형을 데이터 구성과 사용자 특성에 따라 정리하고, 각각의 대표적인 적용 사례를 제시한 것이다.

표 1. 연합학습의 유형

연합학습 유형	데이터 특징	샘플 (사용자)	적용 사례
수평 연합학습	동일 특징	다른 사용자	다국적 병원 간 협력 의료 AI
수직 연합학습	다른 특징	동일 사용자	금융 기관과 전자상거래 플랫폼 간 신용평가
연합 전이 학습	다른 특징 + 다른 사용자	데이터 부족한 기관 간 협력	다국적 기업 간 고객 행동 분석

수평 연합학습은 데이터의 특성은 유사하지만 사용자가 서로 다른 경우에 적용된다. 예를 들어, 여러 병원에서 보유한 환자 데이터는 입력 변수는 유사하지만, 각 병원에 소속된 환자는 중복되지 않으므로 수평 연합학습을 통해 협력 의료 인공지능 시스템을 구현할 수 있다.

수직 연합학습은 서로 다른 특성을 갖지만 동일한 사용자를 보유한 기관 간에 적용된다. 대표적인 사례로는 금융기관과 전자상거래 플랫폼 간의 협업이 있으며, 동일한 고객에 대해 서로 다른 특성의 데이터를 보유하고 있는 기관들이 연합하여 신용 점수를 예측하는 데 활용된다.

연합 전이학습은 사용자와 데이터 특성이 모두 상이한 경우에 적용되며, 주로 데이터가 부족한 기관 간 협업을 통해 모델을 효과적으로 학습할 때 사용된다. 예를 들어, 고객 수가 적은 기업이 다른 기업과 협력해 고객 행동을 분석하는 경우가 이에 해당한다.

다. 연합학습 알고리즘

연합학습은 다양한 목적과 환경에 따라 여러 최적화 알고리즘이 제안되어 왔으며, 각 알고리즘은 모델 업데이트 방식, 클라이언트 간 데이터 이질성 처리, 통신 효율 및 학습 안정성 측면에서 고유한 특성을 가진다.

표 2는 주요 연합학습 알고리즘들의 핵심 방식과 장단점을 정리한 것으로, 알고리즘 선택 시 고려해야 할 다양한 요소들을 비교한 것이다.

표 2. 주요 연합학습 알고리즘 비교

알고리즘	핵심 방법	주요 장점	주요 단점
FedSGD	각 클라이언트의 그래디언트 평균	실시간 학습 가능, 빠른 업데이트	통신량 과다, 민감한 모델 변동성
FedAVG	로컬 모델 파라미터 평균	통신 효율 우수, 빠른 수렴	데이터 이질성(Non-IID)에 취약
FedRep	공통 feature extractor + 개인 head 분리	높은 개인화 성능, 통신 절약	구조 복잡도 증가, 통합 난이도
FedForest	로컬 결정 트리 + 전역 앙상블	계산량 절감, 모델 경량화	확장성 제한, 모델 일반화 어려움
SCAFFOLD	Control Variate를 통한 드리프트 보정	클라이언트 편향 보정, 안정적 수렴	구현 복잡도, 계산량 증가
Fed-SB	LoRA 기반 파라미터 부분 학습	통신량 대폭 감소, 구조 특화 가능	LoRA 적용 제한, 구조 설정 필요
FedRRM	서버 EM 기반 representation 업데이트	이질성 강건성, 빠른 수렴	구현 복잡도, 연산 자원 소모

본 연구에서는 이 중 대표적인 알고리즘인 FedSGD와 FedAVG, 그리고 개인화에 중점을 둔 확장형 알고리즘인 FedRep을 비교 대상으로 선정하였다. FedSGD와 FedAVG는 연합학습에서 가장 기본적인 방식으로, 통신 구조와 연산 효율성 측면에서 성능 비교의 기준선으로 적합하다. 반면 FedRep은 개인화와 데이터 이질성 대응에 강점을 갖춘 최신 기법으로, 향후 멀티태스크 학습 환경으로의 확장을 고려한 분석에 적합하다 [3 - 7].

Ⅲ. 알고리즘 구조 및 성능 비교

가. 비교 대상 알고리즘 및 데이터셋 구성

본 논문에서는 멀티태스크 확장을 고려하여 연합학습 알고리즘의 구조적 특성과 성능을 비교하기 위해, FedSGD, FedAVG, FedRep의 학습 프로세스를 도식화하였다.

그림 2는 FedSGD의 구조를 나타낸 것이다. 각 클라이언트는 로컬 데이터를 기반으로 단일 미니배치를 사용하여 손실함수에 대한 그래디언트를 계산한 후, 이를 서버로 전송한다. 서버는 수신한 그래디언트를 평균하여 글로벌 그래디언트 ∇F(손실 함수 F에 대한 기울기)를 계산하고, 이를 통해 글로벌 모델 파라미터 w를 업데이트한다.

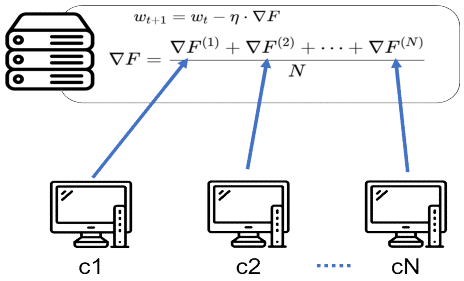

그림 2. FedSGD 구조도

그림 3은 FedAVG의 구조를 보여준다. 각 클라이언트는 로컬 데이터를 기반으로 다수의 에포크(epoch)를 수행하여 모델 가중치 $w^{(k)}$를 업데이트한 후 서버에 전송한다. 서버는 클라이언트로부터 받은 가중치를 평균하여 글로벌 모델을 생성하며, 이를 다시 클라이언트에게 배포한다.

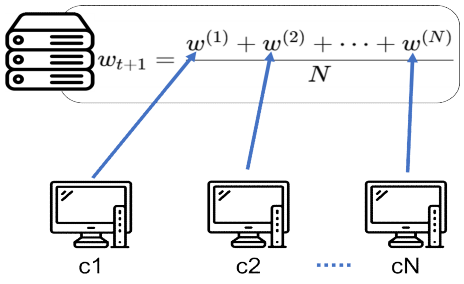

그림 3. FedAVG 구조도

그림 4는 FedRep의 학습 구조를 시각화한 것이다. FedRep은 모델을 공통 feature extractor f와 클라이언트별 개인화 head $h^{(k)}$로 분리하여 학습하며, 각 클라이언트는 feature extractor만 서버에 전송한다. 서버는 이들을 평균하여 새로운 글로벌 feature extractor f_{t+1} 을 생성하고, 다시 클라이언트에 배포하여 다음 라운드를 진행한다. 개인화된 head는 로컬에 유지되어 사용자 맞춤형 학습이 가능하다.

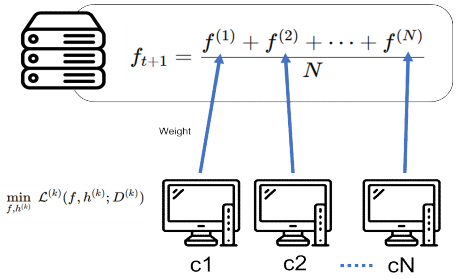

그림 4. FedRep 구조도

이 알고리즘의 성능 비교를 위해 사용된 데이터셋은 FEMNIST 데이터셋이다. 이는 3,500명의 필자가 작성한 62개의 문자(숫자 0~9, 대문자 A~Z, 소문자 a~z)를 포함한 손 글씨 이미지로 구성되며, 각 이미지는 28×28픽셀의 흑백 이미지이다.

이 데이터셋은 사용자의 필기체 특성에 따라 클라이언트 단위로 데이터가 분할되어 있어 Non-IID 환경을 시뮬레이션하는 데 적합하고, 멀티태스크와 유사한 실험 환경 조성이 가능하다. 특히 사용자 중심의 분할 구조는 개인화 성능 측정에 유리하며, 향후 객체 탐지·분할 등의 비전 멀티태스크로의 확장을 고려할 때도 구조적으로 적합하다.

나. 대상 학습 모델

학습 모델은 CNN(Convolutional Neural Network) 기반으로 구성되었으며, 그 구조는 그림 5에 제시되어 있다.

모델은 총 3개의 합성곱 계층(Conv2D)과 2개의 완전연결 계층(Fully Connected layer)으로 이루어져 있으며, 이미지 분류를 위한 표준적인

구조를 따른다.

첫 번째 합성곱 계층은 1개의 입력 채널에서 32개의 필터를 적용하여 특징을 추출하고, ReLU 활성화 함수를 사용한다.

두 번째 합성곱 계층은 32개의 입력 채널에서 64개의 필터를 적용하며, ReLU 함수와 MaxPooling 연산을 통해 주요 특징을 추출한다.

세 번째 합성곱 계층은 64개의 입력 채널에서 128개의 필터를 적용하고, ReLU와 MaxPooling을 다시 적용하여 심화된 특징을 추출한다.

이후 Flatten 레이어를 통해 2차원 출력을 1차원 벡터로 변환하며, 완전연결 계층에서는 128개의 뉴런을 통해 학습이 이루어진다.

마지막 출력층에서는 총 62개의 클래스를 대상으로 분류를 수행하며, 이는 숫자 및 영문 대소문자를 포함하는 FEMNIST 데이터셋의 클래스 수와 대응된다.

해당 모델은 이미지 분류 문제에 적합할 뿐 아니라, 본 논문의 연합학습 실험에서 비교 알고리즘들의 성능을 공정하게 비교하기 위한 공통 기준 모델로 활용되었다.

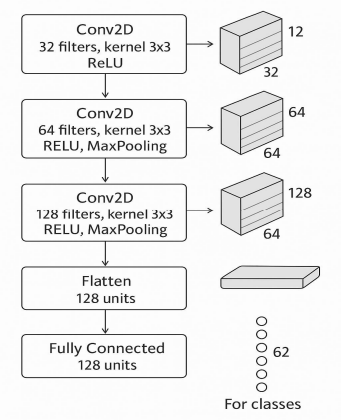

그림 5. 사용한 모델 구조

다. 실험 환경 구성

각 클라이언트는 FEMNIST 데이터셋 상의 서로 다른 사용자의 데이터를 보유하도록 구성되었으며, 이는 클라이언트 간 데이터 분포가 서로 다른 Non-IID 환경을 자연스럽게 형성한다.

또한, 평가 지표는 Loss, 통신 효율성과 Test Accuracy를 기준으로 설정하여 비교 분석한다. 통신 효율성은 70%를 달성하기까지 소요된 라운드 수를 기준으로 정의하였다.

표 3. 실험 환경

항목	설정값
클라이언트	20개
데이터 분포	Non-IID(사용자 기반 분할)
학습 라운드 수	100라운드
평가 지표	Loss, 통신효율성, Test Accuracy

또한 클라이언트 별 샘플 수 및 클래스 분포의 편차는 시각 자료(그림 6, 그림 7)를 통해 확인할 수 있다. FEMNIST는 수평 연합학습(Horizontal FL)에 적합한 데이터셋으로, 사용자의 필기체 스타일을 기반으로 Non-IID 환경으로 클라이언트 수를 20개로 설정하여 실험을 진행하였다.

그림 6은 Non-IID 환경에서 각 클라이언트가 보유한 샘플 수의 분포를 막대그래프로 나타낸 것이다. 대다수 클라이언트는 25,000개 이상 샘플을 보유하고 있는 반면, 일부 클라이언트는 상대적으로 적은 수의 데이터 샘플만을 보유하고 있어 클라이언트 간 데이터 양의 불균형이 존재함을 확인할 수 있다.

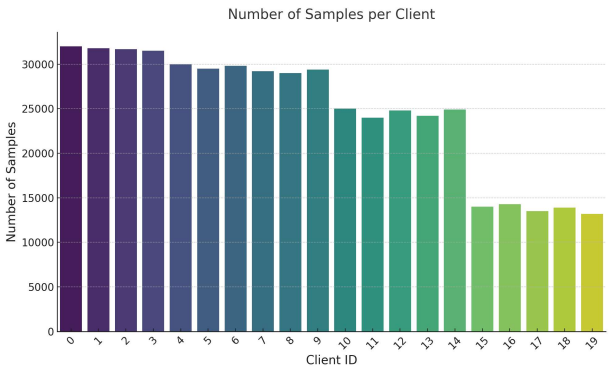

그림 6. 클라이언트 별 샘플 수 분포 그래프

그림 7은 Non-IID 환경으로 구성하여 분할한 FEMNIST 데이터의 클라이언트 별 클래스의 분포도를 도식화한 것이다. 그림 7의 각 행은 클라

이언트를, 각 열은 클래스를 나타내며, 숫자(0~9), 영어 대문자(10~35), 영어 소문자(36~61)로 구성된 62개 클래스를 분류하여 각각 청색, 녹색, 주황색 세 가지 색으로 구분하여 시각화하였다.

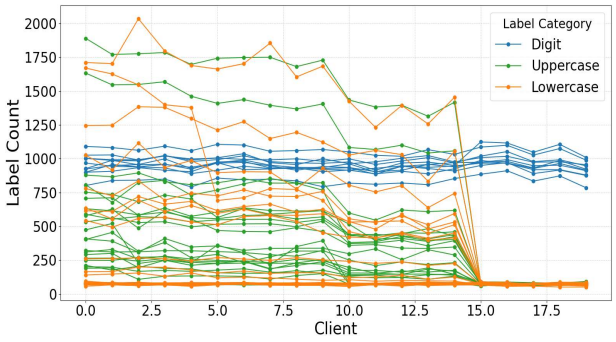

그림 7. 클라이언트 별 클래스 분포 그래프

클라이언트 간 데이터 분포의 비균형은 범주별 샘플 수의 통계에서도 확인할 수 있다. 표 4에 제시된 바와 같이, Digit 범주의 경우 모든 클라이언트에 균일하게 포함되도록 평균 약 95.79개, 표준편차 17.94 수준으로 분포를 설정하였다. 반면, Uppercase와 Lowercase 범주는 각각 평균 82.16개(표준편차 47.76), 69개(표준편차 42.95)로 상대적으로 큰 편차를 가지며, 일부 클라이언트에만 집중되거나 거의 포함되지 않도록 구성한다.

이러한 분포는 클라이언트 간 클래스 구성의 불균형을 유도하며, 연합학습 환경에서의 Non-IID 특성을 효과적으로 반영한다. 특히, 특정 범주의 데이터가 일부 클라이언트에 2,000개 이상 집중되거나, 거의 포함되지 않도록 분할 함으로써, 모델이 학습 과정에서 겪는 분포 차이에 대한 영향을 분석할 수 있도록 한다.

표 4. 클라이언트 간 범주별 분포 통계

범주	평균 샘플 수	표준편차 (Std. Dev.)	분산 (Variance)
Digit(0~9)	95.79개	17.94	321.71
Uppercase(10~35)	82.16개	47.76	2,280.72
Lowercase (36~61)	69개	42.95	1,844.42

라. 실험 결과 및 분석

본 논문에서는 FedSGD, FedAVG, FedRep 세 가지 알고리즘의 성능을 비교하고, 이를 중앙집중형 방식과 함께 평가하였다. 각 알고리즘은 동일한 환경에서 100라운드 동안 학습되었으며, 성능 평가 지표로는 손실(Loss)과 통신 효율성, 정확도(Accuracy)를 사용하였다. 중앙집중형 모델은 전체 데이터를 단일 서버에서 학습하여, 비교 기준선으로 설정하였다.

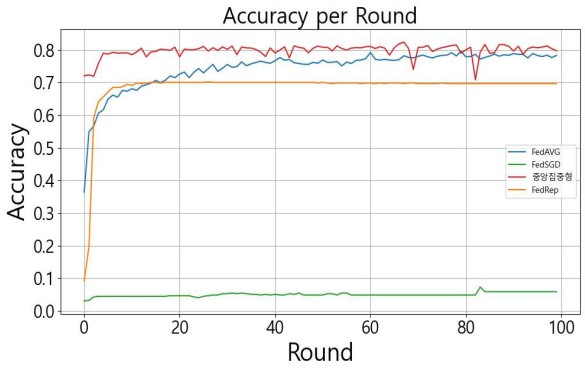

그림 8. 연합학습 알고리즘 별 정확도 변화 비교

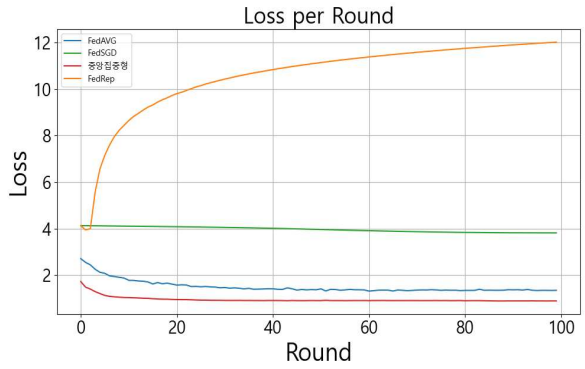

그림 9. 연합학습 알고리즘 별 손실 변화 비교

FedAVG는 100라운드 동안 점진적이며 안정적인 정확도 상승을 보였으며, 최종 정확도는 약 72%에 도달하였다. 손실 또한 일정하게 감소하며 빠르고 안정적인 수렴 특성을 보였다.

FedRep은 매우 빠른 초기 수렴 속도를 기록하여 약 10라운드 이내에 70% 정확도에 도달하였고, 최종 정확도는 78.15% 수준에서 안정적으로 유지되었고, 초기 수렴 속도는 가장 우수하였다.

그러나 손실값은 라운드가 진행됨에 따라 점차 증가하는 경향을 보였고, 이는 클라이언트별 개인화 학습의 결과로 모델 간 편차가 누적되어 나타난 현상으로 해석된다. 이러한 특징은 사용자 맞춤형 서비스에서는 장점이 될 수 있으나, 글로벌 모델의 안정성 측면에서는 한계로 작용할 수 있다.

반면 FedSGD는 학습 전반에 걸쳐 정확도 향상이 거의 발생하지 않았으며, 최종 정확도는 3.67% 수준에 그쳤다. 손실 또한 높은 값을 지속적으로 유지하였다. 이는 FedSGD가 클라이언트별로 단일 스텝의 로컬 업데이트 후, 해당 gradient만을 서버에서 평균화하는 방식으로 인해 클라이언트 간 데이터 분포의 편향을 보정하지 못한 결과로 판단된다. Non-IID 환경에서는 전체 gradient 방향이 왜곡되기 쉬우며, 이는 실제 실험을 통해 확인되었다.

중앙집중형 모델은 전체 데이터셋을 단일 서버에서 학습함으로써 가장 높은 성능을 보였으며, 최종 정확도는 약 80%를 상회하였다. 이는 데이터가 집중된 환경에서 학습되는 경우의 최적 학습 성능을 반영하는 기준선(baseline)으로 해석할 수 있다. 연합학습은 개인정보 보호 및 로컬 학습이라는 장점이 있지만, 정확도 면에서는 중앙집중형 모델에 비해 성능 차이가 존재한다. 향후 연구에서는 이러한 성능 차이를 줄이기 위해 연합학습 기반 모델의 정밀도 향상, Non-IID 데이터 대응 전략, 통신 효율 최적화 등이 필요한 것으로 보인다.

표 5. 최종 결과 지표

지표	FedAVG	FedRep	FedSGD	중앙 집중형
최종 손실	2.181	2.5425	9.0571	0.8866
통신 효율성	35 round	15 round	Not converged	N/A
최종 정확도	71.86%	78.15%	3.67%	82.36%

Ⅳ. 결 론

본 논문은 Non-IID 환경에서 연합학습 알고리즘의 성능을 비교 분석하고, 중앙집중형 학습 방식과의 차이를 실험적으로 평가하였다.

FedAVG는 높은 정확도와 우수한 통신 효율성을 바탕으로, 실제 시스템 적용에 적합한 안정성을 갖춘 알고리즘임을 확인하였고, FedRep은 빠른 수렴과 개인화에 강점을 나타냈으나, 글로벌 손실 값이 증가하는 불안정성이 존재하였다. 반면, FedSGD는 Non-IID 환경에서 낮은 정확도(3.67%)를 기록하며, 실질적인 적용에는 어려움이 큰 것으로 나타났다.

동일한 데이터 조건에서 수행된 중앙집중형 모델은 최종 정확도 82.36%, 손실 0.88 수준으로 가장 뛰어난 성능을 보였으며, 이는 연합학습 알고리즘의 성능이 중앙집중형에 비해 격차가 존재함을 의미하며, Non-IID 환경 대응 및 정확도 향상에 대한 추가적 보완이 필요함을 시사한다.

본 논문은 단일 태스크 기반 실험을 중심으로 진행되었으나, 이는 향후 멀티태스크 비전 문제(예: 분류, 객체 탐지, 의미론적 분할 등)에 연합학습을 적용하는 데 있어 기초 성능 비교 자료로서 활용될 수 있다. 특히, 각 알고리즘의 통신 효율성, 수렴 안정성, 정확도 특성은 멀티태스크 연합학습 알고리즘 설계의 핵심 기준이 될 수 있다.

향후 연구에서는 SCAFFOLD, FedEM 등 Non-IID에 강건한 알고리즘 도입과 함께, 비동기 환경, 클라이언트 참여율 제어, 통신 최적화 전략을 결합하여, 지속적인 지식 교환이 가능한 멀티태스크 연합학습 프레임워크를 설계할 계획이다.

향후 클라이언트 간 데이터 분포의 시간적·구조적 불균형을 정량적으로 분석하기 위해 히든 마르코프 모델(Hidden Markov Model, HMM) 기반의 분포 추적 기법을 연합학습과 연계하는 방법도 고려 중이며, 이는 Non-IID 환경에서 발생하는 데이터 불균형 원인을 진단하고, 개인화 및 알고리즘 맞춤 전략을 설계하는 데 유용한 정보를 제공할 수 있을 것으로 기대된다.

REFERENCES

[1] 이재욱, 서상원, 고한얼, “직/병렬적 학습기반의 연합학습 알고리즘”, *한국소프트웨어종합학술대회 논문집*, 961-963쪽, 제주, 대한민국, 2022년 12월

[2] 오석환. 정송헌. 김경백, “데이터 중독 공격 방어를 위한 신뢰도 점수 기반 연합학습”, *디지털콘텐츠학회논문지*, 제24권 제6호, 1317-1326쪽, 2023년 6월

[3] Lucas Airam C. de Souza, Gabriel Antonio F. Rebello, Gustavo Franco Camilo, Lucas C. B. Guimarães, Otto Carlos M. B. Duarte, “DFedForest: Decentralized Federated Forest”, *2020 IEEE International Conference on Blockchain (Blockchain)*, pp. 90-97, Rio de Janeiro, brazil, 2020.

[4] Liam Collins, Hamed Hassani, Aryan Mokhtari, Sanjay Shakkottai, "Exploiting Shared Representations for Personalized Federated Learning", *Proceedings of the 38th International Conference on Machine Learning (ICML)*, PMLR 139, pp. 2089-2099, 2021.

[5] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank J. Reddi, Sebastian U. Stich, Ananda Theertha Suresh, "SCAFFOLD: Stochastic Controlled Averaging for Federated Learning", *Proceedings of the 37th International Conference on Machine Learning (ICML)*, PMLR 119, pp. 5132-5143, 2020.

[6] “Fed-SB: A Silver Bullet for Extreme Communication Efficiency and Performance in (Private) Federated LoRA Fine-Tuning”(2025),https://arxiv.org/abs/2502.15436, (accessed Feb., 18, 2025).

[7] “FedEM: A Privacy-Preserving Framework for Concurrent Utility Preservation in Federated Learning” (2025), https://arxiv.org/html/2503.06021v1, 2025년 2월 19일

[8] 이주원, 전석환, 방준일, 김화종, “정확도 기반 참여 제한 연합학습 연구”, *한국정보기술학회 논문지*, 제 21권 제 12호, 47-57쪽, 2023년 12월

[9] Taki Hasan Rafi, 임석영, 채동규, “연합학습 기술 및 사업화 동향”, *정보과학학회지*, 제42권 제9호, 21-27쪽, 2024년 9월

[10] 송승엽, 박희재, 박래혁, “IoT 네트워크를 위한 연합학습: 응용 분야 및 한계”, *2024년도 한국통신학회 동계종합학습발표회 논문집*, 232-233쪽, 용평, 대한민국, 2024년 1월

[11] 이재욱, 고한얼, “Non-IID 데이터 환경에서의 정응적 연합학습 기법”, *한국통신학회논문지*, 제49권 제 8호, 1,118-1,120쪽, 2024년 8월

[12] 여웅기, 이준우, “안전한 연합학습 기술 연구 동향 및 분석”, *2024년도 한국통신학회 하계종합학술발표회 논문집*, 563-564쪽, 2024년 6월

[13] 서유리, 조현기, 우하은, 허의남, “어텐션 방법 기반의 멀티모달 감정인식을 위한 프라이버시 보호 Peer-to-Peer 연합학습“, *한국정보과학회 2024 학국소프트웨어종합학술대회 논문집*, 159-161쪽, 2024년 12월

[14] 전석환, 이주원, 방준일, 김화종, ”연합학습 시 글로벌 모델 성능 향상을 위한 기여도 기반 집계 알고리즘“, *한국정보기술학회논문지*, 제21권 제12호 59-66쪽, 2023년 12월

[15] 박혜빈, 박현정, 김낙훈, ”Non-IID 환경에서 연합학습 성능 분석 연구“, *2024년도 대한전자공학회 하계학술대회 논문집*, 3099-3102쪽, 제주, 2024년 6월

저 자 소 개

강태욱(준회원)

2025년 순천대학교 정보통신공학과 학사 졸업

2025년~현재 순천대학교 정보통신공학과 석사 재학

<주관심분야 : 딥러닝, 머신러닝, 데이터분석>

박철영(정회원)

2010년 순천대학교 정보통신공학과 학사 졸업.

2012년 순천대학교 정보통신공학과 석사 졸업.

2017년 순천대학교 정보통신공학과 박사 졸업.

2025년~현재 순천대학교 정보통신공학과 조교수

<주관심분야 : 인공지능 응용시스템, 딥러닝, 머신러닝, 클라우드 컴퓨팅>

이명배(정회원)

2003년 2월 : 순천대학교 컴퓨터공학과(공학사)

2009년 2월 : 순천대학교 컴퓨터과학과(이학석사)

2013년 2월 : 순천대학교 정보통신공학과 박사 수료

2017년 ~ 현재 : 순천대학교 정보통신공학과 연구원

2023년 ~ 현재 : 순천대학교 지능기술연구소 책임연구원

<관심분야 : USN, 빅데이터분석시스템, 농업IT 융합>

신창선(종신회원)

1996년 우석대학교 전산학과 학사 졸업.

1999년 한양대학교 컴퓨터 교육학과 석사 졸업.

2004년 원광대학교 컴퓨터공학과 박사 졸업.

2005년~현재 순천대학교 인공지능공학부 교수

<주관심분야 : 머신러닝, 분산시스템>