# 노드 인지 기반 선택적 학습을 통한 연합학습의 안정성 및 정확성 향상 기법 연구

김유준○ 홍충선*
경희대학교 컴퓨터공학과
{yj4889○, cshong* }@khu.ac.kr

# Node-Aware Selective Learning Based Federated Learning Method for Stability and Accuracy Improvement.

Youjun Kim○, ChoongSeon Hong*
Department of Computer Science and Engineering, Kyung Hee University

**요 약**

최근 다양한 분야에서의 머신러닝(machine learning) 기술 활용도가 높아지고 있다. 또한 사용자 데이터의 프라이버시를 보존할 수 있는 연합학습이 등장하였다. 제한된 통신대역폭, 디바이스의 컴퓨팅 능력은 연합학습에서의 장애가 된다. 따라서 본 논문은 연합학습(federated learning)을 블록체인 기반으로 구현한다. 이는 모델 데이터의 무결성, 투명성 등을 보장하고 사용자의 프라이버시(privacy)를 보호하며 사용자의 참여를 유도할 수 있는 등 많은 강점을 가지고 있다. 이에 더해 네트워크 과부하를 최소화하기 위해 연합학습에 참여하는 사용자 디바이스를 블록체인 노드의 address로써 인지하여 선택하며 로컬 정확도를 모델 통합과정에서 가중치로 적용하여 안정적이고 빠른 모델 학습이 가능하도록 한다.

## 1. 서 론

센서(sensor) 네트워크와 통신 기술의 발전으로 사람과 사람뿐만 아니라 사물과 사물끼리의 통신이 가능해짐에 따라 데이터의 양이 폭발적으로 증가하였다[1]. Cisco의 예측에 따르면 2020년에 500억 개의 디바이스가 인터넷에 연결될 것이라고 한다[2]. 증가한 데이터의 양은 인공지능기술 발전의 원동력이 된다.

인공지능이 급속도로 발전하면서 인공지능 관련 서비스가 점차 다양해지고 있다. 특히 딥러닝(deep learning) 기술이 두각을 나타내고 있다. 딥러닝의 주요 특징 중 하나는 신경망이다. 신경망은 input units, multiple hidden layers, output units 으로 구성되어 있으며 dataset을 input으로 분류, 인식, 예측 등의 결과가 output으로 나오게 된다[3]. 딥러닝은 이러한 신경망으로 이루어져 있으며 이미지 분석, 음성인식, 문자인식 등 다양한 분야에서 성공을 거두고 있으며 캐싱(caching), 자원할당 등 네트워킹 분야에서도 적용되고 있다[4].

전통적인 모델학습은 높은 컴퓨팅 능력을 요구하므로 충분한 컴퓨팅능력을 가진 중앙서버에서 데이터를 수집하고 처리한다. 하지만 대부분의 데이터는 프라이버시(privacy) 문제에 민감하여 모델학습을 위한 데이터 수집에 한계가 있다. 그러나 최근 모바일, Internet of Things(IoT) 등 디바이스의 발전으로 기기는 각 기기 내에서 모델을 학습시킬 수 있을 정도의 컴퓨팅 능력을 갖게 되었다. 이에 따라 [5]는 모델 학습과정에서의 프라이버시(privacy) 문제를 해결할 수 있는 연합학습(federated learning)을 제시하였다.

연합학습은 중앙 서버에서 각 디바이스에 모델을 보내고 각각의 디바이스는 자신의 데이터를 이용하여 디바이스 내에서 모델을 학습시킨다. 학습된 모델은 중앙 서버에 의해 통합된다. 연합학습에 참여하는 사용자의 수는 수 천, 수 만대가 될 수 있다. 그 결과 같은 시간대에 엄청난 수의 사용자가 연합학습에 참여하여 모델을 전송하고 전송받기 때문에 네트워크 과부하가 일어날 수 있다. 네트워크 과부하를 줄이는 가장 효과적인 방법 중 하나는 모델이 목표정확도에 빠르게 수렴하는 것이다. 연합학습의 또 다른 문제점은 사용자로 하여금 자신의 자원을 사용해가며 학습에 참여할 동기가 부족하다는 것이다. 이에 본 논문은 블록체인 기반의 연합학습을 제시한다. 기존의 연합학습은 학습에 참여할 노드를 무작위로 선택하고 오직 데이터의 양만을 고려하여 학습하였다.

이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2019-0-01287, 분산 엣지를 위한 진화형 딥러닝 모델생성 플랫폼) Dr. CS Hong is the corresponding author.

제시한 방법은 연합학습에 참여하는 모든 사용자를 블록체인 노드의 고유 주소로써 구별하고 최적으로 선택하며 로컬 정확도를 연합학습의 가중치로 사용하여 목표 정확도에 빠른 수렴이 가능하게 함으로써 네트워크 부하를 최소화시킨다. 또한 블록체인의 암호 화폐를 인센티브로써 학습 참여자들에게 제공하여 참여 동기를 부여하고 모델 데이터의 무결성, 투명성을 보장한다.

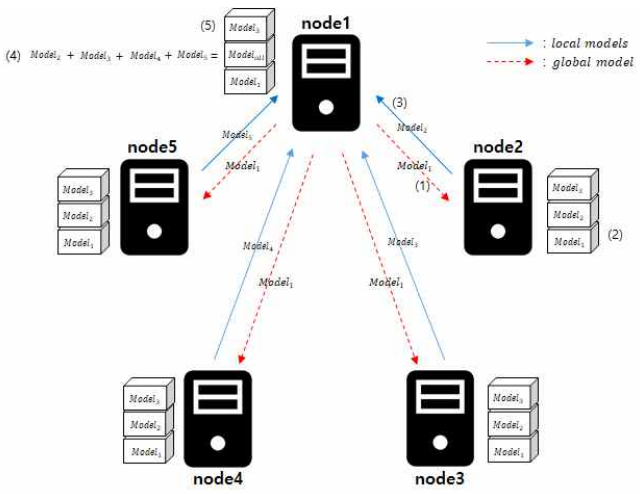
## 2. 관련연구

### 2.1블록체인(blockchain)

블록체인기술은 비트코인과 함께 [6]의 저자에 의해 처음 제시되었다. 모든 거래는 블록으로 저장되며 각 블록은 이전 블록의 해쉬(hash)값을 포함하고 있으므로 서로 연결되어 있다. 연결된 모든 블록은 중앙 서버가 아닌 Peer to Peer(P2P)방식으로 합의 알고리즘을 통해 모든 참여자 디바이스에 저장된다. 이러한 특징으로 블록체인의 거래장부 무 결성, 투명성 등을 보장하여 위, 변조 될 수 없다.

## 3. 제안사항

블록체인 기반 연합학습은 그림 1과 같이 구성되어 있다. 각 node는 모델학습이 가능할 정도의 컴퓨팅 자원을 가지고 있는 말단 디바이스로써 스마트폰, IoT 디바이스 등이 될 수 있다. 모든 node는 채굴자(miner)가 될 수 있다. 합의 알고리즘으로 작업증명(proof of work)[5]을 선택하였다. 그리고 node1을 특정 모델을 학습시키고자 하는 사용자로 가정하였다. 그림 1은 블록체인 기반 연합학습 구상도에 대한 설명이다.




[그림 1] 블록체인 기반 연합학습 구상도

(1) 먼저 node1은 학습시키고자 하는 모델과 로컬 accuracy를 측정하기 위한 test dataset을 학습에 참여한 노드 중 일부에게 전송하기 위해 트랜잭션(transaction)을 발생시킨다.

(2) 발생된 트랜잭션은 채굴자(miner)에 의해 채굴되고 블록화 되어 체인에 저장된다.

(3) 모델은 각 노드의 데이터를 이용하여 각각의 노드에서 학습하고 학습된 모델과 test dataset으로 평가한 accuracy를 node1에게 전송하기 위해 트랜잭션을 발생시킨다.

(4) 발생된 모든 트랜잭션(각 node들에 의해 학습된 각각의 모델)은 채굴자(miner)에 의해 채굴되고 블록화 된다.

(5) node1은 개별 학습된 모든 모델을 로컬 정확도를 가중치로써 모델통합을 한다.

(6) (1) ~ (5)을 원하는 모델 정확도가 나올 때까지 반복한다.

(7) 채굴자(miner)가 채굴을 통해 얻은 인센티브 중 일부는 많은 데이터를 제공한 사용자순으로 적절히 분배한다.

로컬에서 학습된 모든 모델은 블록체인에 저장되고 데이터 무결성 성격을 가진다. 블록체인을 구성하는 블록에는 딥러닝 모델, 모델을 보내는 노드의 주소, 모델을 받은 노드의 주소 등의 데이터로 구성되어 있다. 또한 (7)에 의해 사용자(node)는 학습에 참여한 대가로 인센티브를 얻게 된다.

본 논문에서는 (1)과 (5)에 각각 학습참여노드 선택 방법, 로컬학습 정확도에 따른 가중치 적용방법을 제시한다. 기존의 연합학습과 달리 블록체인환경에서는 노드를 블록체인 주소로써 정확히 인지할 수 있다. 이에 따라 노드의 학습 참가 횟수는 노드의 블록체인 주소로써 node1에서 카운트 된다. 학습참여 노드 선택 방법은 다음과 같다.

표 1. 변수 정의

| Notion | Description |
|---|---|
| $C$ | a fraction of participant nodes |
| $S$ | a fraction of candidate nodes which connected to blockchain. |
| $k_n$ | number of $n$th node' s data set |
| $k$ | number of all node' s data set |
| $w$ | weight of deep learning model |

블록체인에 참여한 전체 노드 중 학습에 가장 조금 참여한 노드들 중 비율 $S$만큼의 노드를 학습참여 후보로써 선택한다. 학습참여 후보 노드 중 비율 $C$만큼의 노드를 선택하여 학습시킨다. 이에 따라 노드의 학습참여 빈도가 어느 한쪽으로 치우치지 않게 된다. 노드들의 공평한

학습참여는 안정적인 학습을 가능하게 하여 학습속도를 향상시킨다. 파라미터는 표 1에 정리되어 있다.

기존 연합학습은 ①식과 같이 모델통합과정에서 각 노드의 데이터 크기만을 고려하였다.

$$w_{t+1} \leftarrow \sum_{n=1}^{N} \frac{k_n}{k} w_{t+1}^{n} \qquad ①$$

하지만 제시된 블록체인 기반 연합 학습은 로컬학습 정확도를 통합모델에 적용한다. 로컬학습 정확도를 측정하기 위해 로컬 노드는 node1에 의해 보내진 test dataset을 이용하고 학습된 모델과 함께 node1에게 보내진다. node1은 받은 정보를 바탕으로 ②식을 활용하여 모델을 통합한다. $a_n$는 학습에 참여한 노드의 n번째 로컬학습 정확도이다. node1에 전송된 모든 로컬모델은 node1의 컴퓨팅 자원을 사용하여 자신이 보유한 train data set에 의해 정확도가 측정된다. 모든 로컬 모델 정확도의 합은 $a = \sum_{n=1}^{N} a_n$ 이다.

$$w_{t+1} \leftarrow \sum_{n=1}^{N} (\frac{a_n}{a}) w_{t+1}^{n} \qquad ②$$

로컬 정확도 가중치 적용방법은 기존 연합학습보다 빠른 학습속도를 보인다.

## 4. 성능평가

본 논문은 학습 데이터 셋으로 MNIST 데이터를 non-iid [5], massively distributed, unbalanced 환경으로 사용하였고 training data 6만개, test data 9천개 로컬 학습 accuracy를 측정하기 위한 데이터 1천개로 구성하였으며 $C$, $S = 0.1$으로 하였다.

그림 2의 그래프는 제시한 방법과 기존 연합학습의 정확도를 비교하여 보여준다. 제시한 연합학습이 fed_avg (기존 연합학습) 보다 accuracy가 더 빠르게 상승하는 것을 볼 수 있다. 얼마나 빠르게 특정 정확도에 도달하는지를 표 2에 나타내었다. 표 2에서 accuracy가 91에 도달하기 위해 fed_avg는 497번의 학습이 필요하며 proposed method는 217번의 학습이 필요한데 이는 제시된 방법의 성능이 2배 이상 더 빠르게 학습한다는 것을 나타낸다. 즉, 각 노드의 학습 참여수를 카운트함으로써 학습 참여율이 낮은 노드를 선택하여 모든 노드가 골고루 학습이 이루어지게 되고 학습의 안정성이 증가하게 된다. 그리고 가중치 적용방법은 학습이 더 잘된 모델에 더 많은 가중치를 주고 통합함으로써 학습의 정확성이 더 높아지는 것을 볼 수 있다. 제안된 방법이 기존 연합학습보다 2배 이상 더 빠른 학습이 가능하며 이는 제한된 대역폭에서 노드 간 통신횟수를 2배 이상 줄여준다는 것을 의미한다.

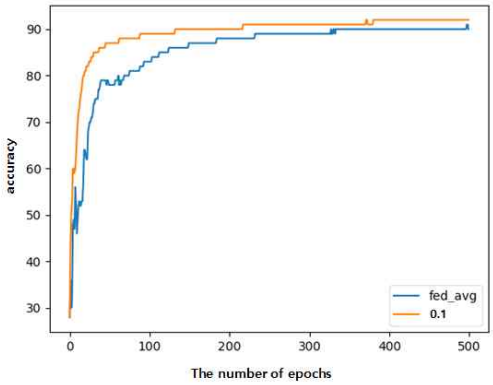



그림 2. 제시된 방법과 federated learning의 학습 당 정확성 비교

표 2. 도달 accuracy round

| accuracy | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 |
|---|---|---|---|---|---|---|---|---|
| fed_avg | 112 | 124 | 149 | 184 | 232 | 327 | 497 | |
| proposed method | 30 | 37 | 45 | 62 | 88 | 132 | 217 | 371 |

## 5. 결론

본 논문은 블록체인 기반의 연합학습 환경에서 모델의 목표 정확도의 빠른 수렴을 위해 학습 안정성 및 정확성을 향상시키기 위해 학습노드 선택, 로컬 모델의 정확도를 가중치로써 적용하는 방법을 제시하였다. 이는 블록체인의 데이터 무 결성, 투명성 등을 보장하고 학습참여에 대한 효율적인 인센티브와 더불어 기존 연합학습이 비해 향상된 성능을 보인다.

## 6. 참고문헌

[1] Qingchen Zhang Laurence T.Yang Zhikui Chen Peng Li, "A survey on deep learning for big data",Information Fusion Volume 42, pages. 146-157, July 2018,

[2] Seok Won Kang, Choong Seon Hong, "Openflow based on Edge Cloud Structure for Efficient Packet Forwarding in Distributed Cloud Environment ",Korea Computer Congress(KCC 2018), page.1306-1308, June 2018

[3] Marcus, G. "Deep learning: A critical appraisal." arXiv:1801.00631[cs.AI], 2018.

[4] Kyi Thar, Nguyen H Tran, Thant Zin Oo, Choong Seon Hong, "DeepMEC: Mobile Edge Caching Using Deep Learning," IEEE Access, Vol.6, Issue 1, pp.78260-78275, December 2018

[5] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data", Proceedings of the 20 th International Conference on Artificial Intelligence and Statistics (AISTATS) 2017. JMLR: W&CP, volume. 54, 2017.

[6] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. [Online] Available: http://www.bitcoin.org/bitcoin.pdf, [downloaded. 01.May.2019]