

# 전기 수요 예측을 위한 연합학습 통신 비용 트레이드 오프

염성웅, 이미리내, 쉬바니, 김경백

전남대학교

yeomsw0421@gmail.com, alflsoquf1207@gmail.com, shivanikolekar@gmail.com,

kyungbaekkim@jnu.ac.kr

## Communication Cost Trade-off in Federated Learning for Electrical Demand Prediction

Sungwoong Yeom, Mirinae Lee, Shivani Sanjay Kolekar, Kyungbaek Kim

Dept. Artificial Intelligence Convergence, Chonnam National University

### 요약

사물 인터넷의 활성화로 인해 빌딩에서의 단기 에너지 소비 예측은 현대의 분산형 전력 시스템에 필수적인 역할을 한다. 사용자의 불규칙한 에너지 소비 패턴에 의한 높은 변동성을 예측하기 위해 시계열 딥러닝 신경망을 사용하는 혁신적인 전기 수요 예측 기법이 등장하였다. 그러나 각 빌딩에서 수집되는 데이터의 양이 다르기 때문에 특정 빌딩에 위치한 모델은 데이터의 양이 부족할 수 있다. 중앙 집중식 시스템은 불균등한 데이터를 가진 빌딩의 모델을 적절하게 학습하기 위해 에너지 소비 샘플 데이터들을 중앙 서버에서 수집하여 모델을 일반화시킨다. 하지만, 이 시스템은 클라이언트에서 서버로 송신되는 민감한 소비 이력 데이터로부터 개인 정보 보호가 보장되지 않을 수 있다. 또한, 이 시스템은 클라이언트의 네트워크 및 에너지 자원이 모두 제한되기 때문에 장기적 관점에서 네트워크의 대역폭에 악영향을 미칠 수 있다. 이 논문에서는 연합학습 프레임워크 환경의 중앙 서버와 클라이언트 사이에서 발생하는 통신 오버헤드와 전기 수요 예측 모델 성능에 대한 트레이드 오프를 연구한다. 제안된 연합학습 프레임워크는 모델 학습 시간 및 개인 정보 보호 측면에서 중앙 집중식 프레임워크를 능가할 수 있음을 보여준다.

### I. 서론

재생 가능한 전력 생산 및 시장의 분산화가 증가함에 따라 빌딩 에너지 관리 시스템에서 지속 가능한 에너지 계획 및 생산을 위한 단기 에너지 소비 예측 연구가 활성화되고 있다 [1]. 하지만 빌딩 거주자들의 수와 이들이 사용하는 IoT(Internet of Things) 장치들의 수가 증가함에 따라 빌딩 내 에너지 수요가 불규칙해지고 있다 [2]. 최근 비선형적인 에너지 수요 예측 성능을 증가시키기 위해 시계열 딥러닝 기법을 사용하는 연구가 활성화되고 있다 [3]. 스마트 빌딩은 시계열 딥러닝 모델의 학습을 위해 구내 고급 계량 인프라(Advanced Metering Infrastructure)를 설치하여 상당한 양의 전기 소비 이력 데이터를 수집하지만 이는 모델을 학습하기에 충분하지 않을 수 있다. 빌딩 관리자는 충분하지 않은 데이터 양을 소유한 빌딩들을 위해 에너지 소비 이력 데이터를 다수의 빌딩들로부터 클라우드로 집계하는 중앙 집중식 접근 방식을 채택해왔다 [4][8]. 하지만, 수백만 개의 스마트 미터에서 수집되는 데이터가 모델 학습에 사용되는 경우 모델을 일반화하기 어렵고 계산 집약적이며 확장이 어려울 수 있다. 또한, 수백만 개의 스마트 미터로부터 수집되는 데이터에 의해 발생하는 대역폭으로 인해 상당한 통신 비용이 든다.

에지 컴퓨팅은 IoT와 같은 엔드 서비스 유저에게 보다 더 가까운 원격지 컴퓨터에 계산을 오프로딩(Offloading)하기 위해 도입된 분산 패러다임으로 클라우드 컴퓨팅을 네트워크 에지로 확장하여 속도, 효율성, 안정성, 개인정보 보호, 보안 및 확장성 측면에서 다양한 이점을 제공한다 [5]. 에지 컴퓨팅은 다수의 빌딩들의 AMI로부터 데이터 수집 및 확장할 때 발생하는 개인 정보 보호 및 확장성 문제를 효율적으로 극복하기 위한 훌륭한 솔루션이다. 그러나 딥러닝 기반 에너지 수요 예측 모델은 여러 빌딩 거주자로부터 데이터를 중앙 서버에서 수집하여 학습되는 경우에만 모델

의 일반화를 보장하기 때문에 이는 데이터들이 분산된 상태로 유지되고 로컬에서 처리되어야 하는 엣지 컴퓨팅의 사례와 대조된다.

이러한 딥러닝과 엣지 컴퓨팅 사이의 간극을 해소하기 위해 연합학습이 제안되었다 [6]. 연합학습은 중앙 엔티티의 조정하에 모델 학습에 참여하는 장치들의 연합하에 글로벌 모델을 학습하는 분산 머신러닝 기법이다. 중앙 서버는 집계된 모델의 가중치를 통해 글로벌 딥러닝 모델(예: 전기 수요 예측을 위한 시계열 딥러닝 신경망)을 초기화하고 K개의 클라이언트와 공유한다. 즉, 중앙 서버는 각 클라이언트(빌딩 거주자 또는 빌딩)에게 글로벌 모델의 사본을 호스팅 한다. 이후 각 클라이언트는 로컬에서 모델을 학습시킨다. 이때, 로컬에 호스팅 되는 딥러닝 모델은 자체 데이터 세트를 사용한다. 이제 클라이언트는 원시 데이터를 송신하는 대신 로컬에서 업데이트된 모델의 가중치를 중앙 서버로 송신한다. 중앙 서버는 집계 단계에서 수신된 모델들의 가중치를 FedAVG[7]를 통해 평균화하여 글로벌 모델을 업데이트한다. 이 프로세스는 글로벌 모델의 손실이 수렴될 때까지 여러 라운드에 걸쳐 학습을 수행한다.

연합학습은 원시 데이터 대신 로컬에서 업데이트된 모델의 가중치를 송신함으로써 개인 정보를 보호하고 클라이언트 통신 비용을 줄일 수 있지만, 연합학습에 참여하는 클라이언트의 수가 증가한다면 모델의 가중치를 송수신하기 위해 발생하는 상당한 양의 링크 대역폭을 소비하기 때문에 통신 비용과 글로벌 모델의 최종 손실 값 사이의 균형을 완전히 이해하는 데 여전히 격차 있다. 또한, 링크 대역폭이 유한하기 때문에 연합학습은 각 라운드 수를 조정함으로써 리소스 제약을 설정해야 한다.

이 논문에서는 스마트 빌딩 환경을 위한 클라이언트에서 서버로 로컬 업데이트된 매개변수를 공유할 때 발생하는 통신 비용과 글로벌 모델을 최종 테스트 손실 간의 균형을 맞추기 위한 트레이드 오프를 제안한다. 우

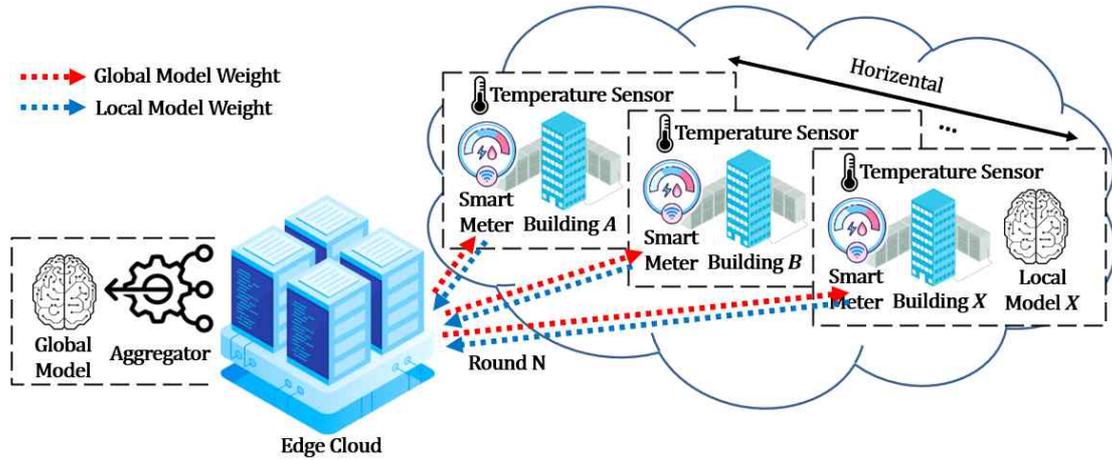


그림 1. 빌딩 에너지 수요 예측을 위한 연합학습 아키텍처

리는 연합학습에 참여하는 클라이언트 수, 모델의 정확도, 학습 시간 간의 균형을 고려하여 다양한 실험을 수행한다.

## II. 빌딩 에너지 수요 예측을 위한 연합학습 프레임워크

그림 1은 제안하는 빌딩 에너지 소비 예측을 위한 연합학습 프레임워크를 보여준다. 제안하는 아키텍처는 시계열 딥러닝 신경망 중 하나인 LSTM(Long Short-Term Memory) 모델을 각 빌딩에 위치시키고 이를 분산학습하여 빌딩의 에너지 소비를 정확하게 예측한다. 제안하는 아키텍처는 제로 에너지 빌딩과 엣지 클라우드 2가지로 구성된다. 제로 에너지 빌딩은 태양열 패널 등을 통해 재생에너지를 생산 및 저장하고 빌딩 거주자들에게 에너지를 공급해준다. 제로 에너지 빌딩은 거주자들의 다양한 IoT 장치들을 통해 소비한 전기 요금을 빌딩 거주자들에게 청구할 수 있다. 이때, 전기 소비 패턴을 예측한다면 제로 에너지 빌딩에 저장된 에너지와 비교하여 부족한 에너지를 스마트 그리드로부터 가져올 수 있다. 이에 따라 제로 에너지 빌딩은 스마트 미터 및 빌딩 에너지 예측 시스템 두 가지로 구성된다. 스마트 미터 및 IoT 장치들은 빌딩 구내에 설치되어 단위 시간당 소비되는 전기 이력 데이터 및 기후 데이터를 측정 및 기록한다. 빌딩 에너지 예측 시스템은 수집한 데이터를 통해 LSTM 모델을 학습한다. 에지 클라우드는 제로 에너지 빌딩들을 관리하는 중앙 서버로서 여러 빌딩 에너지 예측 시스템에서 학습한 LSTM 모델의 파라미터들을 집계 및 학습 후 배포한다.

### 1. LSTM 기반 빌딩 에너지 수요 예측 기법

빌딩 에너지 수요 예측 성능의 중요한 점은 빌딩 에너지 수요 패턴에서 복잡 계절성이 나타나는 것이다. 제안된 접근 방식은 빌딩 에너지 수요의 복잡 계절성을 학습하기 위해 전기 수요, 실내 온도 및 상대 습도와 같은 다변량을 분해한 후 에너지 수요를 추정하는 LSTM을 사용한다. 그림 2는 제안된 LSTM 기반 전기 수요 예측 모델 아키텍처를 나타낸다. 이 모델은 3개의 LSTM 레이어와 3개의 Dense 레이어로 구성된다. 이 세 개의 Dense 계층 중 처음 두 계층은 ReLU(Rectified Linear Unit) 활성화 함수를 사용하고 마지막 계층은 선형 활성화 함수를 사용한다. 각 레이어는 10개의 노드를 포함한다. 에너지 수요는 하루 동안 일정한 시간 간격으로 관찰된다. 관찰된 에너지 수요에 대한 예측의 안정성을 보장하기 위해 제안된 LSTM 모델에 대한 입력 벡터가 생성된다. 입력 벡터는 시별로 관찰되는 전기 수요, 실내 온도, 실내 습도, 월별 인덱스, 시별 인덱스, 요일 유형 세 가지 특성으로 구성된다.

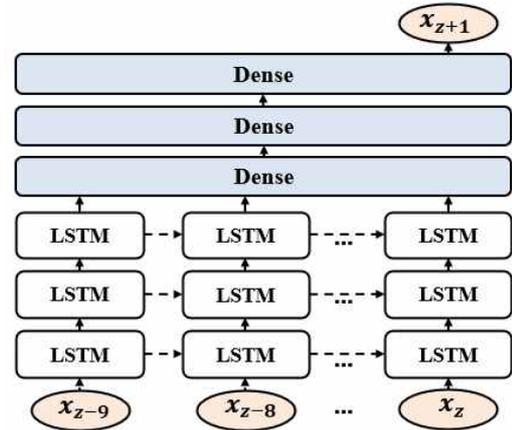


그림 2. LSTM 기반 빌딩 에너지 수요 예측 모델 아키텍처

제안하는 LSTM 모델은 전기 수요를 예측하기 위해 입력 벡터를 학습한다. 구체적으로 제안하는 LSTM 모델은 특정 시간부터 9시간 전까지 10개의 연속된 에너지 수요 정보를 사용하여 다음 1시간 후에 대한 에너지 수요를 예측한다.

제안하는 모델의 입력 벡터는 6가지 도메인으로 구성된다. 입력 벡터의 첫 번째, 두 번째 및 세 번째 도메인은 전기 수요, 실내 온도 및 상대 습도이다. 각 입력 벡터의 도메인은 사람의 행동에 영향을 받는다. 예를 들어, 사람의 행동은 오전에 점진적으로 활동이 증가하며 오후에는 활발히 활동하고 저녁에는 점차 활동이 감소한다. 이에 따라 빌딩은 사람에게 쾌적한 실내의 온도 및 습도를 맞추기 위해 에너지를 사용할 수 있다. 즉, 각 도메인은 일별, 주별, 월별 계절성이 섞인 복합 계절성이 나타난다. 이 복합 계절성은 단기 전기 예측에 큰 영향을 미칠 수 있다. 우리는 전기 수요, 실내 온도 및 상대 습도로부터 STL 분해 기법[11]을 사용하여 각 도메인 별로 계절성, 추세, 잔차를 분해하여 이를 입력 벡터로 사용한다. 입력 벡터의 네 번째 및 다섯 번째 도메인은 월별 및 시별 도메인이다. 월별 인덱스는 1에서 12 사이의 값을 갖는다. 시별 인덱스는 1에서 24 사이의 값을 갖는다. 마지막 입력 벡터의 여섯 번째 도메인은 요일 유형이다. 요일 유형은 공휴일을 포함하여 요일에 따라 달라지는 도메인들의 특성들을 표현하기 위해 8가지의 범주를 고려하여 원-핫 인코딩을 통해 임베딩한다. 따라서, 전기 수요 예측을 위한 LSTM 모델의 총 입력 벡터 길이는 19이다. 제안하는 LSTM 모델은 빌딩 에너지 관리 시스템에 설치된다.

## 2. 빌딩 에너지 수요 예측을 위한 FedAvg

연합학습은 각 빌딩의 데이터 샘플 수와 관계없이 모델을 공동으로 학습할 수 있도록 제안되었다. 예지 클라우드는 글로벌 모델을 초기화하고 초기화된 글로벌 모델의 가중치  $w_0^g$ 를 연합학습에 참여하는 빌딩들  $S = \{s_1, \dots, s_n\}$ 에게 전송한다.  $i$ 번째 빌딩  $s_i$ 에서는 로컬에서 모델을 학습하기 위해 학습률  $\eta$  및 배치 크기  $b$ 와 같은 하이퍼 파라미터를 초기화하고 전달받은 모델 가중치  $w_0^{s_i}$ 를 업데이트한다. 라운드  $t$ 에서 학습률  $\eta$  및 배치 크기  $b$ 를 사용하여 학습한  $i$ 번째 빌딩의 모델 가중치 업데이트에 대한 수식은 아래와 같다.

$$w_t^{s_i} = w_t^g - \eta \nabla \text{loss}(w; b) \quad (1)$$

학습된 모델 가중치  $w_t^{s_i}$ 는 예지 클라우드로 전송된다. 예지 클라우드에서는  $n$ 개의 빌딩들로부터 업데이트된 가중치들  $w_t^{s_1}, \dots, w_t^{s_n}$ 을 집계하고 이를 FedAvg 알고리즘[7]을 통해 학습한다.

$$w_t^g = \sum_{s_i \in S} \frac{|x_{s_i}|}{\sum_{s_j \in S} |x_{s_j}|} w_t^{s_i} \quad (2)$$

이때,  $x_{s_i}$ 는  $i$ 번째 빌딩의 데이터 셋 크기를 의미하며 글로벌 모델의 가중치  $w_t^g$ 를 FedAvg 사용하여 평균화한다.

## III. 검증

### 1. 실험 설정

이 장에서는 제안된 빌딩 수요 예측을 위한 연합학습(Federated Learning)의 성능을 검증하기 위해 중앙 집중식 학습(Centralized Learning)을 비교한다. 우리는 건물 에너지 시뮬레이터인 CitySim[0]에 의해 미리 계산된 CityLearn[10]의 빌딩 전기 수요 데이터 셋을 실험에 사용한다. CityLearn 환경은 4개의 익명화된 기후 구역과 기후 구역별로 9개의 건물 클러스터로 구성되어 있다. 각 빌딩 데이터 셋은 1년 동안 1시간 단위로 측정된 전기 수요 데이터로 구성된다. 우리는 이중 첫 번째 기후 구역에 위치한 7개의 빌딩 데이터 셋을 사용하여 테스트베드 실험을 수행한다. 이 실험은 2개의 NVIDIA GeForce RTX 2080 SUPER가 설치된 머신에서 도커 컨테이너를 기반으로 1개의 MongoDB 서버, 1개의 Chief 노드, 그리고 7개의 Worker 노드로 구성된 클러스터를 통해 수행된다. 표 1은 각 빌딩 데이터에 대한 평균 전기 수요, 평균 실내온도, 평균 상대 습도를 보여준다. 이 빌딩의 전기 수요는 실내 습도 및 온도에 따라 달라질 수 있으며 빌딩별로 위치한 LSTM 모델은 서로 다른 기후 및 전기 수요의 패턴을 학습하여 연합학습에 영향을 미칠 수 있다. 각 데이터 셋의 첫 10개월은 LSTM 모델을 학습하는데 사용하고 다음 2개월은 테스트를 위해 사용한다. LSTM 모델 학습을 위해 Adam Optimizer와 크기가 70인 배치를 사용하였다. 우리는 RMSE(Root Mean Square Error)를 사용하여 연합학습과 중앙 집중식 학습의 성능을 비교하였다.

### 2. 평가결과

그림 5는 중앙 집중식 학습에서 주기별로 전송되는 데이터에 의해 소요되는 파일 크기와 연합학습에서 LSTM 모델의 가중치가 송수신에 소요되는 파일 크기에 대한 비교를 보여준다. 이때, 이 실험은 단일 Worker 노드에서 수행되었다. 중앙 집중식 학습에서 1년 동안 1시간 단위로

	Avg. Equipment Electric Power [kWh]	Avg. Indoor Temperature [C]	Avg. Relative Humidity [%]
Building_1	24.83	23.34	46.66
Building_2	24.89	40.81	8.83
Building_3	23.09	54.85	6.42
Building_4	23.01	58.22	4.22
Building_5	23.13	52.32	11.34
Building_6	22.54	52.29	10.83
Building_7	23.67	50.46	11.49

표 1. 빌딩별 전기 수요 데이터 셋

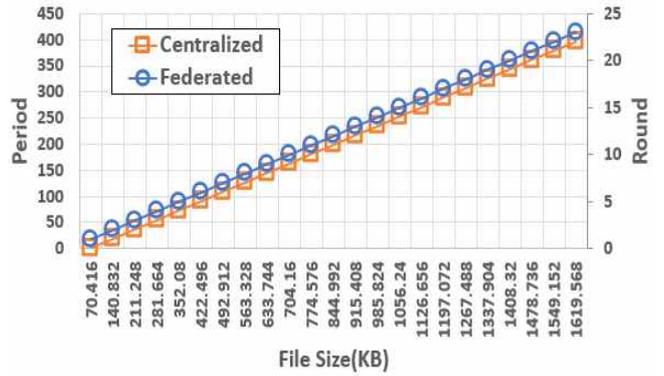


그림 5. 중앙 집중식 학습의 주기별로 전송되는 전기 수요 데이터에 대한 파일 크기와 연합학습의 라운드별로 공유되는 LSTM 모델의 가중치에 대한 파일 크기 비교

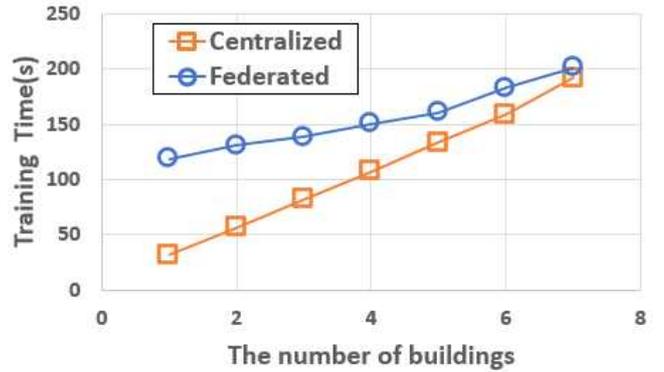


그림 6. 빌딩 수에 따른 중앙 집중식 학습과 연합학습에 소요되는 학습 시간 비교

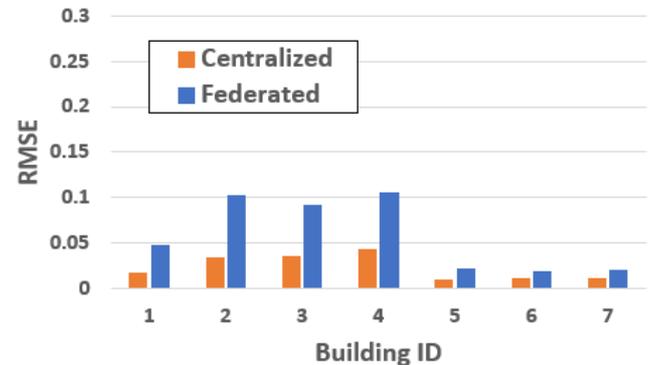


그림 7. 빌딩 에너지 수요 예측을 위한 중앙 집중식 학습 모델과 연합학습 모델의 빌딩별 RMSE 비교

클라이언트에서 중앙 서버로 보내지는 데이터의 총 양은 약 1478KB이고 이는 연합학습에서 약 20 라운드 동안 소요되는 데이터의 총 양과 같다. 따라서, 이 실험에서 연합학습의 라운드는 20으로 설정한다. 하지만, 중앙 집중식 학습과 연합학습은 참여하는 Worker 노드의 개수가 증가함에 따라 학습 시간이 길어질 수 있다.

그림 6은 각 모델 학습에 참여하는 빌딩의 수가 증가함에 따라 소요되는 학습 시간을 비교한다. 연합학습에 대한 소요시간은 참여하는 Worker 노드의 수가 적을수록 중앙 집중식 학습에 대한 소요시간에 비해 크다. 하지만 빌딩 수에 따른 연합학습에 대한 소요시간의 증가율은 중앙 집중식 학습에 대한 소요시간의 증가율에 비해 적기 때문에 빌딩의 수가 증가할수록 연합학습의 총 소요시간은 중앙 집중식 학습의 총 소요시간에 비해 적다. 제안하는 실험환경에서는 학습에 참여하는 빌딩의 수가 7개 일 때 두 그래프가 교차한다. 따라서, 이 실험에서 연합학습에 참여하는 Worker 노드 수는 7개로 설정한다. 우리는 이러한 환경에서 중앙 집중식 학습의 RMSE와 연합학습의 RMSE를 통해 성능을 비교한다.

그림 7은 빌딩 에너지 수요 예측을 위해 학습된 중앙 집중식 학습 모델과 연합학습 모델의 RMSE를 비교한다. 중앙 집중식 학습 모델은 학습에 참여한 빌딩의 기후 및 전기 수요 데이터들을 집계하고 다양한 패턴을 학습하기 때문에 각 빌딩에서 낮은 RMSE를 보여준다. 반면에 연합학습 모델은 로컬에서 학습된 모델들의 가중치를 평균화하기 때문에 다른 패턴의 데이터를 학습한 모델의 가중치가 연합학습에 참여한다면 일부 빌딩에 위치한 모델의 RMSE가 높아진다. 빌딩 1,2,3,4에서 연합학습 모델의 RMSE는 중앙 집중식 학습 모델의 RMSE가 낮다. 하지만, 빌딩 5,6,7에서 연합학습 모델의 RMSE와 중앙 집중식 학습 모델의 RMSE는 비슷하다. 즉, 빌딩 1,2,3,4의 데이터 소비 패턴과 빌딩 5,6,7의 데이터 소비 패턴이 다를 수 있다. 이 두 군집은 서로의 학습을 방해할 수 있다.

## V. 결론

이 논문에서 빌딩에서 효과적인 단기 전기 수요 예측을 위해 연합학습 프레임워크를 제안하였다. 우리는 제안된 프레임워크의 성능을 평가하기 위해 수집된 방대한 양의 데이터를 사용하여 LSTM 모델이 중앙 서버에서 학습되는 중앙 집중식 프레임워크와 비교하였다. 결과적으로 제안된 연합학습 프레임워크가 중앙 집중식 학습 프레임워크와 부분적으로 비슷한 성능을 보여주지만 모델 학습 시간 및 개인 정보 보호 측면에서 중앙 집중식 학습 프레임워크를 능가할 수 있음을 보여준다. 더욱이 전기 소비 이력이 1시간 이하의 단위로 측정한다면 통신 오버헤드 또한 줄일 수 있음을 보여준다. 하지만, 다른 소비 패턴을 가진 빌딩 데이터를 학습한 모델의 가중치를 연합학습에 참여시키면 모델의 성능이 낮아질 수 있다. 차후에 전기 수요 예측을 위한 연합학습의 성능을 향상시키기 위한 연합 클러스터링에 대한 연구를 진행하고자 한다.

## ACKNOWLEDGMENT

이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(IITP-2019-0-01343).

## 참 고 문 헌

[1] A. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan and N. A. Khan, "Load forecasting dynamic pricing and DSM in smart grid: A review", *Renew. Sustain. Energy Rev.*, vol. 54, pp. 1311-1322, Feb. 2016.

[2] Y. Peng, Y. Wang, X. Lu, H. Li, D. Shi, Z. Wang, et al., "Short-term load forecasting at different aggregation levels with predictability analysis", *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT Asia)*, pp. 3385-3390, May 2019.

[3] Wang, Jian Qi, Yu Du, and Jing Wang. "LSTM based long-term energy consumption prediction with periodicity." *Energy* 197 (2020): 117197.

[4] P. Zhang, X. Wu, X. Wang and S. Bi, "Short-term load forecasting based on big data technologies", *CSEE J. Power Energy Syst.*, vol. 1, no. 3, pp. 59-67, Sep. 2015.

[5] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob and A. Ahmed, "Edge computing: A survey", *Future Gener. Comput. Syst.*, vol. 97, pp. 219-235, Aug. 2019.

[6] B. McMahan, E. Moore, D. Ramage, S. Hampson and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data", *Proc. Int. Conf. Artif. Intell. Statist.*, pp. 1273-1282, 2017.

[7] Li, Xiang, et al. "On the convergence of fedavg on non-iid data." *arXiv preprint arXiv:1907.02189* (2019).

[8] Global Smart Meters Market—Growth Trends Forecasts (2020–2025), Jan. 2021, [online] Available: <https://tinyurl.com/y4ujl8f3>

[9] Vázquez-Canteli, José R., et al. "CityLearn v1. 0: An OpenAI gym environment for demand response with deep reinforcement learning." *Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*. 2019.

[10] Robinson, Darren, et al. "CitySim: Comprehensive micro-simulation of resource flows for sustainable urban planning." *Proceedings of the Eleventh International IBPSA Conference*. No. CONF. 2009.

[11] Cleveland, Robert B., et al. "STL: A seasonal-trend decomposition." *J. Off. Stat* 6.1 (1990): 3-73.