

LSTM 기반 스택킹 앙상블 기법을 활용한 전력망 내의 전기자동차 충전 전력 수요량 예측 알고리즘

양창석, 안드레스, 김기선*
광주과학기술원

{lovingy92, andrescn, kskim}@gist.ac.kr

LSTM-based Stacking Ensemble for Electric Vehicle Charging Demand Prediction in Power Grid

Changseok Yang, Lismer Andres Caceres Najarro, Kiseon Kim*
Gwangju Institute of Science and Technology.

요 약

최근 여러 기계학습 알고리즘과 딥러닝 신경망을 활용하여 전력망 내의 전기자동차 충전 전력 수요량을 예측하는 연구가 진행되고 있다. 하지만 전기자동차 사용자별 다양한 충전 행동 패턴과 시간대별로 다른 전기자동차 충전 요청 수요의 높은 변동성은 예측 성능을 저하하는 문제를 야기시킨다. 본 연구에서는 기계학습 알고리즘들과 장단기 메모리 순환 신경망(Long Short-Term Memory)을 결합한 새로운 스택킹 앙상블 방법을 제안하였다. Caltech 대학 주차장의 EV 충전 사용 데이터 기반 모의 실험 결과, 제안한 방법이 기계학습 알고리즘과 순환 신경망 방법과 비교하여 RMSE와 MAE 두가지 성능 평가 지표에서 충전 전력 수요량 예측 오차를 개선함을 확인하였다.

I. 서 론

전기자동차 충전은 시간에 따른 충전 부하 변화율이 높기 때문에 다수의 전기자동차(Electric Vehicle, EV)들을 충전할 경우 전력망의 부하 계통에 커다란 문제를 야기할 수 있다. 따라서 전력망 입장에서 실제 데이터 기반 EV 충전 전력량 예측 시스템은 전력망의 안정적인 전력 운용 관리 및 공급 안정화에 있어 매우 중요한 문제이다 [1]. 최근 기본적인 기계학습 알고리즘에서 나아가 앙상블 기법과 딥러닝 신경망을 활용한 전기자동차 충전 전력 수요량 예측을 높이려는 연구가 활발히 진행되고 있으나 서로 다른 기계학습과 딥러닝 알고리즘들을 결합하는 앙상블에 대한 연구는 초기 단계이다.

EV 충전 전력 수요 관련 특성(충전 세션별 사용자 수, 충전 간격 번호, 충전 시간, 요일 정보)를 여러 기계학습 알고리즘들을 통해 학습하여 예측하였다 [2]. 기본적인 특성만 고려한 것에서 나아가 EV 사용자별 과거 충전 행동 패턴을 K-Means와 K-NN 알고리즘을 사용하여 클러스터링 분류하여 이를 개별 기계학습 알고리즘의 추가적인 특성 정보로 활용하여 [2]의 알고리즘보다 예측 성능을 향상한 연구가 있었다 [3]. 하지만 EV 전력 충전은 시간대별 EV 충전 부하율이 다른 특성을 보이는데 [3]의 기계학습 알고리즘은 이러한 시간적 연속성을 고려하지 못한 단점이 있다. 이를 해결하기 위해 기계학습 알고리즘에서 나아가 순환 신경망을 활용하여 EV 충전 전력 수요량을 예측하려는 연구가 진행되었다 [4].

유전학 분야에서 딥러닝 기반의 스택킹 앙상블 방법이 최근 제안된 바 있다. 신경 근육 장애 판별 정확도를 높이기 위해 다중 기계학습 알고리즘들과 컨볼루션 심층 신경망을 결합하여 판별의 정확도를 높였다 [5]. 본 연구는 이에 착안하여 기계학습 알고리즘들과 순환 신경망(Recurrent Neural Network, RNN), 장단기 메모리 순환 신경망(Long Short-Term Memory, LSTM)을 결합한 새로운 딥러닝 신경망 기반의 스택킹 앙상블 구조를 제안하였다. 제안한 방법이 기계학습 알고리즘과 순환 신

경망과의 모의 실험을 통해 EV 충전 전력 수요량 예측 성능을 비교한다.

II. 본 론

EV 사용자별 다른 충전 행동 패턴과 전력망 내의 시간대별 다른 EV 충전 전력 수요량의 변동으로 인해 시간대별 연속성을 고려한 모델 일반화는 과적합(overfitting)을 방지하기 위한 중요한 이슈이다. 본 연구는 시계열 데이터 예측에 높은 성능을 보이는 LSTM 신경망과 모델 일반화에 장점을 보이는 스택킹 앙상블을 결합한 새로운 방법을 제안한다. 본론에서는 스택킹 앙상블 방법과 LSTM 신경망에 대해 설명하고 이를 결합한 시스템을 제안한다.

A. 스택킹 앙상블

스택킹 앙상블은 예측 모델이 과적합(overfitting)으로부터 방지하기 위한 일반화 성능을 향상시키기 위한 목적으로 사용된 앙상블 기법 중 하나이다 [6]. 이 기법은 1단계에서 제안한 개별 회귀 모델들의 일반화 오차를 줄이는 데 효과적인 방법이다. 일반화의 핵심은 2단계에서 메타 모델이 개별 회귀 모델의 예측값들로 이루어진 2차 공간에서 다시 학습, 일반화를 수행함으로써 문제 공간에 대한 더 많은 정보를 얻을 수 있고 단일 모델의 편향(bias) 값을 크게 줄일 수 있다. 또한 배경, 부스팅 방법보다 다른 종류의 단일 모델들을 혼합한다는 점에서 일반화 오차를 줄임으로써 예측 성능을 크게 높이는 데 강점이 있다. 그림 1은 스택킹 앙상블 방법의 구조이다.

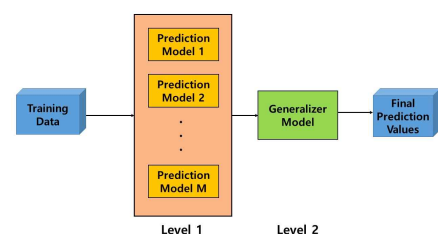


그림 1. 스택킹 앙상블 구조

B. 장단기 메모리 순환 신경망

전력망 내의 전기자동차 충전 전력 수요량은 시간대별로 다른 전기자동차의 충전 요청 수으로 급격한 변동을 보인다. 이러한 점에서 기계학습 알고리즘을 메타 모델로 사용하는 기존의 스택킹 앙상블 방법에서 나아가 본 논문에서는 기계학습 알고리즘이 아닌 시간적 연속 특성을 반영한 LSTM 신경망을 최종 메타 모델로 고안하였다. LSTM은 시계열 데이터와 같은 순차적인 정보를 처리하는 순환 신경망(RNN)에서 나아가 정보 흐름을 제어하는 입력, 출력, 망각 게이트를 사용하여 장기 의존성을 학습할 수 있다. 이를 통해 역전파 학습 과정에서 입력층으로 갈수록 가중치들이 업데이트가 되지 않는 기울기 소실 문제를 해결할 수 있다.

C. 시스템 개요

본 연구에서 제안한 EV 충전 전력 수요량 예측 시스템은 기계학습 알고리즘들이 입력 데이터를 각자 학습하는 1단계와 LSTM 메타 모델이 스택킹한 메타 데이터를 다시 학습하여 최종 예측하는 2단계로 구성된다. 앞서 소개한 [3]의 연구에서는 EV 충전 사용 raw 데이터를 입력 특성(충전 세션별 사용자 수, 충전 간격 번호, 충전 시간, 요일 정보와 사용자별 충전 행동 클래스(5개))들로 가공하였다. 그 후 기계학습 알고리즘(Gradient Boosting, GB)이 특성들을 입력 받아 학습한 후 예측값을 출력하였다. [4]의 연구에서는 데이터를 가공한 후 기본 LSTM 3층 신경망으로 예측값을 출력하였다 (그림 2a 참조). 반면 제안한 시스템에서는 먼저 1단계에서 4개의 기계학습 알고리즘이 [3]의 특성들을 학습하여 개별 예측값을 출력하였다. 다음 2단계에서 순차적 데이터 예측 성능에 강점을 보이는 LSTM 신경망을 메타 모델(generalizer)로 활용하여 메타 데이터(1단계 개별 모델에서 나온 예측값들의 집합)들을 토대로 재학습하여 최종 EV 충전 전력 수요 예측 값을 출력하였다. 그림 2b는 EV 충전 전력 수요량을 예측하는 LSTM 기반의 스택킹 앙상블 알고리즘 시스템의 전체 구조이다.

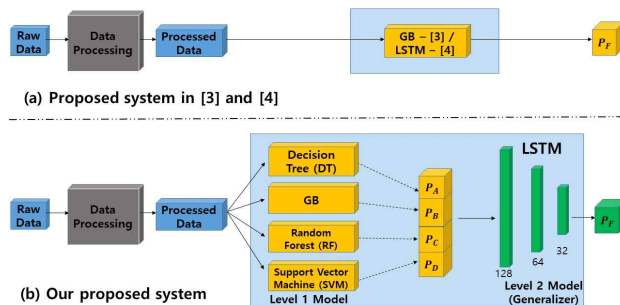


그림 2. 제안한 EV 충전 전력 예측 시스템과 논문 [3],[4]의 시스템 비교

III. 시뮬레이션 결과

본 논문에서는 Caltech 대학 주차장 내의 EV 충전 공공 데이터 셋을 사용하였다 [7]. 전력망의 15분 단위 EV 전력 충전 수요량을 예측하기 위해 11,384개의 학습 데이터(2018.11.1~2019.10.31)와 3,087개의 테스트 데이터(2019.11.1~2020.1.31)로 나눈 후 15분 단위로 샘플링하였다. 사용한 변수는 해당 간격마다 충전을 요청한 사용자들의 수, 충전 시작 시점과 요일, 충전 소요 시간, 사용자들의 충전 행동 패턴을 군집화하여 원-핫 인코딩한 해당 클래스 정보들이다. 1단계에서 10-교차 검증(cross validation)을 통해 튜닝한 4가지의 기계학습 알고리즘(DT, GB, RF, SVM)을 사용해 예측값을 구하였다. 그리고 이러한 예측값들을 묶어 메타 데이터 셋을 만들어 2단계에서 LSTM으로 메타 데이터를 학습하여 최종

예측 값을 구하였다. 학습에 사용한 입력 데이터의 파라미터를 살펴보면 batch size는 일주일 단위인 672, time step은 하루 단위인 96, 특성은 4개의 개별 기계학습 알고리즘 예측값들을 결합한 4개로 구성되어 있다. 네트워크 층의 노드는 128, 64, 32로 3개의 층으로 구성되어 있으며 사용한 활성화 함수는 ReLU이고, epoch은 20이다.

예측 성능 평가 기준은 평균 절대 오차(MAE)와 평균 제곱근 오차(RMSE)이다 :

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \quad (2)$$

N은 학습 데이터 수, y_i 는 실제 전력량과 y'_i 는 예측 전력량이다.

표 1. 제안한 딥러닝 기반 스택킹 앙상블과 다른 기계학습/딥러닝 알고리즘들의 성능 비교 (S.E : Stacking Ensemble)

Algorithm (Generalizer)	Accuracy (kWh)	
	MAE	RMSE
GB [3]	5.34	6.65
LSTM [4]	4.34	6.22
S.E (RNN)	4.12	6.05
S.E (LSTM)	3.88	5.82

표 1은 전력망의 EV 충전 전력 수요량 예측 시스템에서 제안한 알고리즘과 [3]에서 가장 높은 성능을 나타낸 기계학습 알고리즘(GB), LSTM[4]과의 예측 성능을 비교한 것이다. 모의실험 결과 제안한 LSTM 신경망 기반 스택킹 앙상블 방법은 [3]은 기계학습 알고리즘 특성상 EV 충전 전력의 시간대별 주기성을 반영하지 못한 한계로 인해 [4]보다 성능이 낮은 것을 보였다. 또한 기계학습 알고리즘을 결합한 LSTM 기반 스택킹 앙상블 네트워크 구조를 설계 적용한 제안 시스템은 [3]의 EV 사용자별 충전 행동 패턴 클래스 특성과 [4]의 시간대별 연속성을 반영함은 물론 스택킹 앙상블의 장점인 일반화 오차를 줄임으로써 MAE 3.88, RMSE 5.82를 기록하며 두가지 성능 평가 지표 모두 예측 성능이 개선될 수 있었다.

IV. 결론

본 논문에서는 전력망의 15분 단위 EV 충전 전력 수요량 예측 성능을 높이기 위해, LSTM 기반 스택킹 앙상블 알고리즘을 제안하였다. 4개의 기계학습 알고리즘으로 예측을 진행한 후 최종 generalizer가 메타 데이터 셋을 토대로 다시 학습하여 최종 예측함으로써 일반화 오차를 줄였다. 또한 시간대별로 차이를 보이는 충전 전력 수요량의 변동을 반영하기 위해 장단기 메모리 순환 신경망을 최종 generalizer 모델로 설계하여 시간적 연속성과 주기성을 고려할 수 있었다. 그 결과 다른 방법에 비해 제안한 시스템이 예측 성능이 더 높음을 확인하였다.

ACKNOWLEDGMENT

This work was supported by the project Development of Automatic Identification Monitoring System for Fishing Gears funded by the Ministry of Oceans and Fisheries, South Korea.

참고 문헌

- [1] A. Mehrabi et al., "Decentralized Greedy-Based Algorithm for Smart Energy Management in Plug-in Electric Vehicle Energy Distribution Systems", *IEEE Access*, vol. 8, 2020.
- [2] Y. Chung et al., "Ensemble Machine Learning-based Algorithm for Electric Vehicle User Behavior Pattern", *Applied Energy*, vol. 254, 2019.
- [3] C. Yang et al., "Machine Learning-based Charging Demand Prediction in Power Grid Using Charging Behavior Patterns of Electric Vehicle Users", in *Proc. KICS conference*, 2020.
- [4] J. Zhu et al., "A Novel LSTM based Deep Learning Approach for Multi-time Scale Electric Vehicles Charging Load Prediction", in *Proc. IEEE PES Innovative Smart Grid Technologies Asia*, 2019.
- [5] A. Khamaria et al., "A Novel Deep Learning-based Multi-Model Ensemble Method for the Prediction of Neuromuscular Disorders", *Neural Computing and Applications*, vol. 32, 2018.
- [6] D. H. Wolpert, "Stacked Generalization", *Neural Networks*, vol. 5, 1992.
- [7] Z. J. Lee, T. Li, and S. H. Low, "ACN-Data: Analysis and Applications of an Open EV Charging Dataset," in *Proc. the Tenth ACM International Conference on Future Energy Systems*, vol.19, 2019.