

태양광 발전량의 주요변수 분석과 다양한 예측모델 성능 비교

유다송, 이영재*

안동대학교, *한국전자통신연구원

gjffod0826@naver.com, *lyj4295@etri.re.kr

Performance Comparison of Prediction Model and Analysis of Major Feature for Solar Power Generation System

Yu Da Song, Lee Young Jae*

Andong National Univ., *ETRI

요 약

최근 기후변화의 심각성으로 탄소중립과 같은 환경 보호 정책이 시행되면서, 재생에너지 발전설비의 비중이 증가하고 있다.

본 논문에서는 신재생 에너지의 주요 에너지원인 태양광 발전을 예측하기 위해 다양한 모델을 설계해보았다.

머신러닝 모델과 딥러닝 기반의 순환신경망 모델, 총 6가지 예측 모델을 생성하여,
태양광 발전량에 영향을 미치는 주요변수를 분석하고 상대 평가 지표를 활용해 성능을 비교하였다.

I. 서 론

전 세계적으로 기후변화의 심각성을 인지하여, 탄소 순 배출량 제로를 목표로 하는 탄소중립을 선언하고 이를 실현하기 위해 이행체계를 구축하고 있다. 우리나라에서도 2020년 12월에 '2050 탄소중립 추진전략'을 발표하여, 화석연료 중심의 에너지 체계에서 신재생 에너지로 전환하기 위한 전략을 제시하였다[1]. 여기서 핵심 재생 에너지 기술로 태양광과 풍력 발전이 주목받고 있으며, 이를 이용한 발전설비의 비중이 늘어나는 추세이다. 한국 전력 거래소의 전력통계정보시스템에 따르면 2022년 9월 기준 신재생 에너지의 발전설비 비중이 전체의 20.1%를 차지했고 이 중에서 태양광 발전설비가 75%를 달성했다[2]. 이로써 신재생 에너지의 주요 에너지원인 태양광 발전의 역할이 중요해지면서, 안정적으로 전력을 공급하기 위하여 발전량 예측에 관한 연구가 필요하다.

태양광 발전량 예측 연구에는 기후 정보를 독립변수로 사용하여 발전량을 예측하는 연구가 활발히 진행되고 있으며, 예측 모델은 대부분 머신러닝과 딥러닝 기술이 활용되고 있다. 또한, 최근에는 LSTM(Long Short-Term Memory), GRU(Gated Recurrent Units)와 같은 순환신경망(Recurrent Neural Network, RNN)을 활용한 방법이 발전량 예측에 적용되고 있다[3].

본 논문에서는 기상청에서 제공하는 실제 기상관측 자료를 사용해 태양광 발전량 예측 모델을 개발하였다. 머신러닝의 주요 모델과 순환신경망 계열 모델, 총 6가지를 성능 평가 지표를 활용해, 실제 태양광 발전량과 예측된 태양광 발전량의 차이를 비교하고자 한다.

II. 본론

II-1. 실험 데이터

발전량 예측을 위한 데이터는 첫 번째로 한국농어촌공사에서 제공하는 전라남도 진도군의 태양광 발전소 현황 데이터를 사용하였다. 해당 데이터는 2017년 1월 1일부터 2020년 9월까지 1시간 단위로 수집된 태양광 발

전량 자료이다. 두 번째는 발전소 인근의 기상관측 자료로, 진도군에는 자료가 없어 가장 가까운 전라남도 목포시의 기온(°C), 강수량(mm), 풍속(m/s), 습도(%), 일조시간(hr), 일사량(MJ/m²), 전운량(10분위)의 데이터를 활용하였다.

II-2. 데이터 전처리

태양광 발전기의 유지보수 혹은 통신 이상의 문제가 발생할 경우, 발전량이 결측되는 경우가 존재한다. 따라서 진도지역 발전소의 32,856개의 데이터 중에서 24시간 단위를 기준으로 발전량의 합계가 0일 때, 해당 날짜를 제거하여 결측치를 처리하였다. 또한, 기상관측 자료의 변수 단위를 일정한 값에 맞추기 위하여 StandardScaler 방식을 사용해 각 데이터의 평균을 0, 분산을 1로 변환해주었다.

II-3. 학습 모델 선정

본 논문에서 사용한 예측 모델은 6가지로 머신러닝 기법인 Random Forest(RF), Support Vector Machine(SVM), Gradient Boosting Machine(GBM), Extreme Gradient Boosting(XGBoost)과 RNN 방식인 LSTM, GRU를 이용한다. 해당 모델을 이용하여 2017년 1월 1일부터 2019년 12월 31일의 데이터는 모두 학습을 위한 Train으로 사용하였고 2020년 1월 1일부터 2020년 9월 30일까지는 Test로 예측 성능을 평가하였다.

II-4. 연구 결과

특성 중요도는 발전량을 예측하는 과정에서 각 특성이 예측에 얼마나 큰 영향을 미쳤는지를 알려주는 지표이다. Fig 1은 XGBoost와 GBM의 특성 중요도이다. XGBoost에서의 특성 중요도는 일사량이 0.89로 가장 큰 비중을 차지하고 있으며, GBM은 일조시간과 일사량이 함께 사용되었다. 태양광 발전의 특성상 태양의 빛 에너지를 이용하기 때문에 기상 매개 변수 중 일조시간과 일사량이 유의미한 변수로 사용되었다고 볼 수 있다.

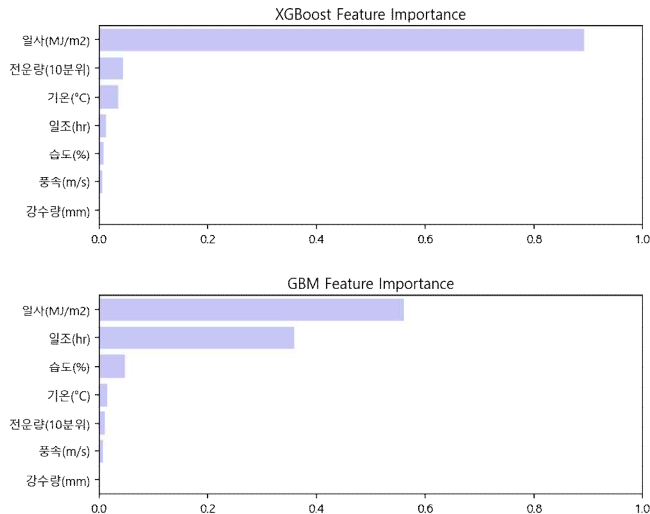


Fig 1. XGBoost와 GBM의 특성 중요도

성능 평가를 위해 실제 발전량과 예측된 발전량의 차이를 비교하고자, 상대 평가 지표인 MAE(Mean Absolute Error), RMSE(Root Mean Squared Error), R-squared Score(Coefficient of Determination)를 사용하였다. 6가지 모델의 성능 평가 지표는 Table 1과 같다.

실험 결과, 머신러닝 모델에서는 XGBoost 모델이, RNN 방식에는 GRU 모델이 높은 성능을 보였다. 전체 예측 모델을 비교하였을 때, 딥러닝의 RNN 방식인 GRU가 37.165의 MAE와 73.505의 RMSE와 0.911의 R-squared Score로 가장 우수한 성능을 보였다. 또한, 기존에 수행되었던 태양광 발전의 연구인 Table 2와 비교했을 때도 좋은 결과를 보인다.

Model	MAE	RMSE	R ²
RF	39.917	78.365	0.879
SVM	40.475	78.286	0.879
GBM	39.646	77.048	0.883
XGB	41.712	77.380	0.882
LSTM	38.175	75.821	0.902
GRU	37.165	73.505	0.911

Table 1. 모델의 성능 평가 지표

Model	Outcome	Literature
ANN	MAE=48.42, RMSE=72.37	Lee et al.,[4]
LSTM	MAE=39.40, RMSE=78.25	Kim et al.,[5]
HS	RMSE=90.0, R ² =0.84	Cho et al.,[6]

Table 2. 기존 태양광 발전의 연구 결과

Fig 2는 성능 평가 지표가 우수하였던 머신러닝 방식의 GBM과 딥러닝 방식의 GRU를 사용하여 2020년 3월 1일부터 2020년 4월 30일까지의 발전량을 실제값과 예측값을 비교해 그래프로 나타낸 것이다. 초록색 선(GBM Predicted value)으로 표시된 GBM의 예측 결과와 비교하여, 주황색 선(GRU Predicted value)으로 표시된 GRU의 예측값이 실제 변동 추이와 매우 유사한 것을 확인할 수 있다. 이는 발전량 예측을 수행함에 있어서, 머신러닝 모델을 사용하는 것보다 딥러닝 기반의 RNN 모델을 사용하는 것이 더 적합한 것을 알 수 있다.

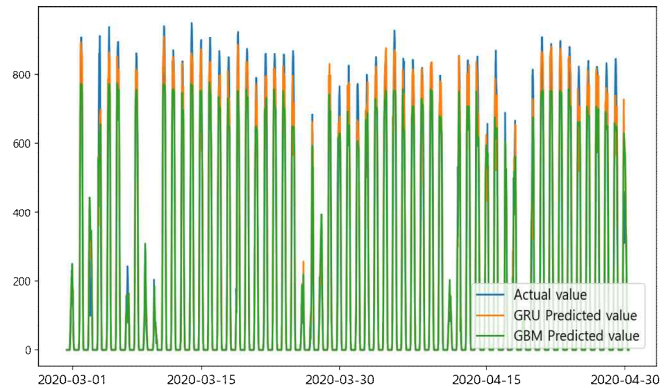


Fig 2. GBM과 GRU를 이용한 발전량 예측 비교

III. 결론

본 논문에서는 태양광 발전량 예측을 위해 진도군의 태양광 발전소의 현황 데이터와 기상 관측 자료를 활용하여 예측 모델을 설계하였다. 모델의 특성 중요도를 통해 기상 매개 변수 중 태양과 연관성이 있는 일조시간과 일사량이 유의미한 변수로 사용된 것을 확인하였다. 또한, 머신러닝 기반의 모델과 딥러닝 기반의 모델을 비교한 결과, 딥러닝 방식 GRU 모델의 예측 결과가 가장 우수한 성능을 보였다.

향후 연구 과제는 현재의 태양광 예측 모델을 발전시킬 예정이다. 현재 이용하고 있는 예측 모델은 기후적 특성으로만 발전량을 예측하고 있어, 계절적 특성을 고려하지 못한다. 따라서 이후 연구에는 계절적 특성을 반영한 모델을 구현해보고 1시간 뒤, 혹은 다음 날의 태양광 발전량을 예측하는 연구를 진행하고자 한다.

ACKNOWLEDGMENT

본 논문은 안동지역 ‘중소기업 ICT 융합기술 경쟁력 강화 사업 (23AD1100)’의 연구결과로 수행되었음

참 고 문 헌

- [1] S. Ha, J. Choi and S. Oh, “Technology Competitiveness Analysis of New & Renewable Energy in Major Countries,” New & Renewable Energy, pp. 72–84, 2022.
- [2] “Korea Electric Power Statistics Information System,” 2022, (<https://epsis.kpx.or.kr/epsisnew/>)
- [3] O. Kwon, S. Choi, H. Jo and H. Cha, “The Prediction of a Floating Photovoltaic Generation Utilizing RNN,” The Transactions of the Korean Institute of Electrical Engineers, pp. 1126~1134, 2022.
- [4] G. Lee and S. Song, “Predicting Solar Power Generation based on Deep Neural Network,” The Korea Contents Association, pp. 375–376, 2022.
- [5] B. Kim, J. Hwang and D. Suh, “A Study on the Short-term PV Generation Prediction using Sliding Window Method,” The Korean Institute of Communication Sciences Conference, pp. 523–524, 2020.
- [6] Y. Cho, G. Lee, J. Hong, and Z. Geem, “Prediction of Photovoltaic Generation Using Machine Learning Models with Various Weight Optimization Techniques,” Journal of the Korean Institute of Intelligent Systems, pp. 1–6, 2022.