

머신러닝 기반 트리 모델을 이용한 태양광 발전량 예측

김예림¹, 손석우¹

¹서울대학교 지구환경과학부

태양광 발전량을 정확하게 예측하는 기술은 빠르게 변화하는 재생 에너지의 환경에서 즉각적인 전력 수요를 효율적으로 충족시킬 뿐 아니라 장기적인 태양광 발전소의 인프라 계획에도 도움이 된다. 최근 연구들은 심층 신경망(Deep Neural Networks)이나 순환 신경망(Recurrent Neural Networks)을 이용하였지만, 이러한 모델들은 모델 구조와 학습 자료 분포에 따라 성능이 크게 좌우된다. 특히 대한민국과 같이 연간 강수 데이터 중 90%가 비어있는 경우 강수가 발생할 때의 발전량에 대한 예측 성능이 떨어지는 경향이 있다. 따라서 본 연구에서는 입력 변수 조합이나 결측값이 많은 변수에 덜 민감한 머신 러닝 기반 부스팅 모델을 사용하여 태양광 발전량을 예측했다. 2013년부터 2020년도의 한국의 영암 F1 태양광 발전소와 기상 관측소에서 측정한 기온, 풍속 그리고 일사 변수를 학습 데이터로 사용하고 2021년부터 2022년의 데이터를 이용하여 예측 성능을 평가하였다. 본 모델과 Liu, Laibao, et al. (2023)의 모델의 예측 성능을 비교한 결과, 트리 기반 모델을 사용하여 예측한 발전량은 기존 모델보다 15% 적은 오차를 보이는 것으로 나타났다. 날짜와 시간 변수, 강수량, 이슬점온도, 일조, 적설 그리고 전운량과 함께 모델을 학습시켰을 때, 오차가 기존 부스팅 모델보다 약 20% 더 개선되었다. 추후에 어떤 변수 조합이 태양광 발전량 예측에 적합한지 분석하여, 빠른 학습시간 안에서 높고 합리적인 정확도를 가진 모델을 개발하고자 한다.

Key words: Gradient Boosting Model, Solar Power Forecasting, CatBoost Model