

기계학습 기법을 조합한 태양광 발전량 예측

김상훈, 윤정미, 권기웅

전자부품연구원

{ksh7150, yunjm, kiwoong.kwon}@keti.re.kr

Photovoltaic power forecasting by combining multiple machine learning techniques

Sanghun Kim, Jung-mi Yoon, Kiwoong Kwon

Korea Electronics Technology Institute

요약

친환경 에너지에 대한 수요에 발맞추어 다양한 전력 생산 기술들이 연구 개발되고 있다. 여러 장점으로 인해, 태양광 발전이 주목받고 있지만, 날씨 변화에 민감하다는 단점이 있다. 따라서 기상 상태에 따른 발전량을 예측하여 활용한다면, 태양광 산업에 도움을 줄 수 있다. 본 논문에서는 몇 가지 기계학습 기법을 조합하여, 태양광 발전량 예측 성능을 개선하고자 하였다. 개별적으로 학습된 몇 가지 예측 결과를 재결합하여, 최종 예측량을 산출하였다. 최종 결과물은 단일 기계학습 기법 대비 성능 개선을 보였다.

I. 서론

산업 발전 및 소득 수준 향상에 따라, 에너지 수요는 꾸준히 증가하고 있다. 특히 전력에 대한 수요가 폭발적으로 증가하고 있으며, 이러한 전력 수요에 부응하기 위하여 다양한 발전 기술들이 연구되고 있다. 특히 환경 오염, 지구 온난화 등의 문제를 해결하기 위하여 친환경, 무공해 에너지 기술이 주목받고 있다 [1].

이러한 친환경 에너지 중에서, 태양광 (Photovoltaic) 발전은 무한 제공되는 태양 에너지를 이용하여 연료가 필요 없고, 대기오염이나 폐기물을 발생시키지 않는다. 또한, 풍력, 파력, 지열 발전처럼 장소에 구애받지 않으며, 비교적 소규모에서 대규모 발전 설비까지 사용자의 요구에 맞추어 설치할 수 있다는 장점이 있다. 그리고, 태양광 발전은 시설 관리가 용이하고, 설비의 수명이 길어 발전 사업자의 편의 및 사업성 측면에서도 주목받고 있다.

그러나 이러한 태양광 발전은, 일사량 변동에 따른 발전량 편차가 매우 크고, 날씨가 좋지 않을 때는 발전할 수 없다는 단점이 있다. 따라서, 발전량을 예측하여 날씨 변화에 따른 대안을 마련하는 것이 매우 중요하다. 이러한 요구에 의해 태양광 발전량 예측을 위한 다양한 연구들이 수행되고 있으며 [2]-[4], 산업계에서도 많은 관심을 가지고 있다.

본 논문에서는 몇 가지 기계학습 기법을 이용하여 태양광 발전량 예측을 수행하였으며, 이를 조합하여 더 나은 결과를 얻도록 노력하였다. 제안한 기법은 개별적인 예측 기술 대비 정확도 개선의 장점을 가지며, 다른 분야로의 확장 가능성을 보였다.

II. 제안한 기법

본 논문에서는 몇 가지 기계학습 기법을 이용하여, 태양광 발전량 예측을 수행하였으며, 개별 기법의 예측치를 조합하여 성능 개선을 시도하였다. 따라서, 본 논문에서 제안한 기법은 앙상블 학습법의 (Ensemble method) 일종이라고 할 수 있다 [5].

먼저 발전량 예측에 사용될 데이터는 기상청에서 제공하는 날씨 예보를 사용하였으며, 기온, 습도, 운량, 강수 형태로 구성되어 있다. 여기에 계절적 특성을 포함하기 위하여, 날짜 및 시간을 추가하여 학습 데이터를 구성

하였다.

발전량 데이터를 수집하기 위한 태양광 패널을 성남시 분당구 야탑동에 설치하였으며, 실시간 발전량 정보를 Database (DB)에 저장하였다. 이러한 데이터는 전처리 과정을 통하여, 이상치 (outlier) 제거, 보간법을 (interpolation) 이용한 손실 데이터 복구, 범위 조절 (scaling) 등을 수행하였다. 평균 및 분산을 통한 이상치 검출을 수행하였으며, 제거 후 cubic spline 보간을 이용하여, 해당 데이터를 복원하였다. 또한 손실된 데이터에 대하여도 동일한 보간법을 수행하였다. 이후 해당 데이터들은 평균 0, 표준편차 1이 되도록 scaling하여 사용하였다.

이후 3종류의 기계학습 기법을 이용하여, 개별 학습을 수행 하였다. 개별 기법은 Deep learning 2종 및 Support vector regression (SVR) 이다. Deep learning 기법 2종류는 내부 구조 및 파라미터 설정을 다르게 하여, 서로 다른 예측 결과를 도출하게 하였다. 이러한 기법들의 성능을 비교하여 표 1에 나타내었다. 세 가지 기법 모두 대등한 성능을 보였으며, 검증 데이터에 따라 성능 순위가 바뀔 수 있어 특정 기법이 우수하다고 단정 짓기에 어려움이 있다.

따라서, 본 논문에서는 개별 기법의 결과를 결합하여 더욱 우수한 성능을 도출하고자 하였다. 전처리 이후의 데이터들은 개별 기법에 학습되며, 이후 실제 발전량과의 차이를 기록하였다. 이러한 오차를 바탕으로 해당 조건에서 가장 정확한 기법을 판별하였으며, 이를 훈련 데이터로 재 사용하였다. 따라서, 기상 변화 및 시간에 따른 다양한 조건 하에서 가장 정확할 것 같은 기법을 찾아내도록 학습시킨 것이다. 마지막 출력단에 소프트맥스 (Softmax) 함수를 이용하여, 확률을 얻었다. 3가지 결과 중 확률이 가장 높은 것을 선택하는 (Most likely)와 개별 확률을 가중치를 사용하여 결과를 결합하는 (Weighted average) 기법을 제안하였다.

이러한 일련의 과정은 그림 1에 나타내었으며, 사용된 데이터는 표 2에 정리하였다.

표 1 개별 기법의 성능 비교

	DNN 1	DNN 2	SVR
우세 빈도 (%)	33.12	32.65	34.23

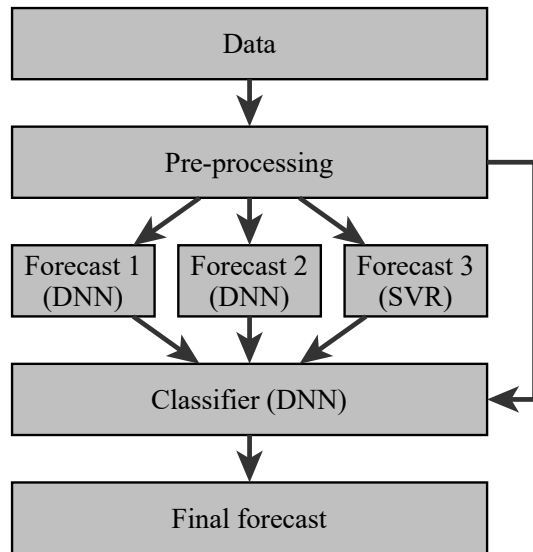


그림 1 시스템 구성도

표 2 학습에 사용된 데이터

훈련 데이터 기간	2018/11/16 ~ 2020/05/31
검증 데이터 기간	2020/06/01 ~ 2020/06/30
사용 데이터	온도, 습도, 운량, 강수 형태, 발전량, 날짜, 시간
데이터 수집 장소	성남시 분당구 야탑동

III. 성능 비교

그림 2는 개별 기법들의 발전량 예측치와 실제 발전량을 비교한 것이다. 개별 기법들은 유사한 성능을 보이지만, 각 구간별로 성능 우위가 뒤바뀔 수 있다. 따라서, 개별 예측치를 잘 결합하는 것만으로도 더욱 우수한 성능을 얻을 수 있을 것이다.

그림 3은 각 기법들의 오차율을 빈도로 나타낸 것이다. 본 논문에서 제안한 기법들이 개별 기법보다 더 작은 오차율을 가질 확률이 높음을 알 수 있다. 따라서, 단일 기법을 사용할 때 보다 다양한 기계학습 기법의 결합을 이용하면, 성능 개선을 얻을 수 있음을 확인하였다.

IV. 결론 및 향후 연구

본 논문에서는 몇 가지 기계학습 기법을 조합하여, 태양광 발전량 예측 성능을 향상 시키고자 하였다. 제안한 기법은 개별 기계학습 기법의 예측량 중 가장 가능성이 높은 것을 선택하거나, 개별 기법의 가능성에 따라 가중치를 부여하여, 성능 개선을 시도하였다.

성능 검증을 통하여, 제안한 기법이 개별 기법 보다 성능이 우수함을 알 수 있었다. 다만 개별 기법의 특성을 파악하고, 조합하는 방법에 관한 추가 연구가 필요하다. 따라서 본 논문의 저자들은 기술 개선을 위해 다양한 시도를 수행할 예정이다.

ACKNOWLEDGMENT

본 연구는 산업통상자원부의 재원으로 한국에너지기술연구원(KETEP)의 지원을 받아 수행한 연구 과제입니다 (No. 2017010000780).

참 고 문 헌

- [1] 유성현, 안춘기, "신재생 에너지 기술 : 태양광 기술과 기후예측 기술과의 융합," *대한전기학회 전기의 세계*, vol. 67, no. 2, pp. 19-25, Feb. 2018.
- [2] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *Plos one*, pp. 1-26, Mar. 2018.
- [3] M. P. Almeida, M. Muñoz, I. de la Parra, and O. Perpiñán, "Comparative study of PV power forecast using parametric and nonparametric PV models," *Solar Energy*, vol. 155, pp. 854-866, Oct. 2017.
- [4] 엄지영, 최형진, 조수환, "기상정보를 활용한 도시규모-EMS용 태양광 발전량 예측모델," *대한전기학회 논문지*, vol. 64, no. 3, pp. 393-398, Mar. 2015.
- [5] C. M. Bishop, *Pattern recognition and machine learning*, NY: Springer, 2006.

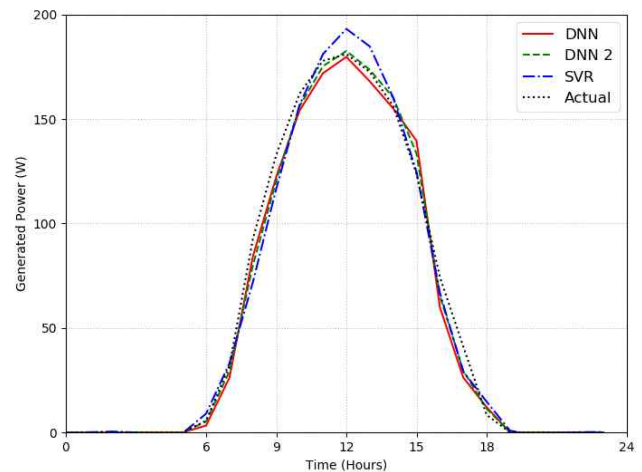


그림 2 개별 기법들의 발전량 예측

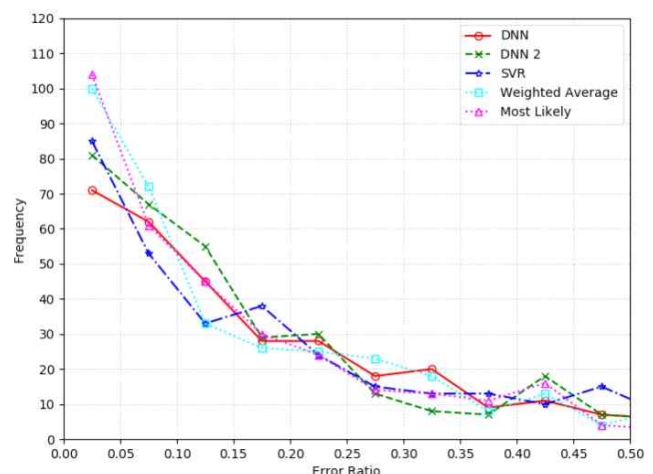


그림 3 개별 기법과 새로운 기법의 성능비교