

태양광 에너지 예측을 위한 SVM 및 ANN 모델의 성능 비교

정원석¹ · 정영화¹ · 박문규² · 이창교³ · 서정욱^{1*}

¹남서울대학교 · ²세종대학교 · ³구미전자정보기술원

Performance comparison of SVM and ANN models for solar energy prediction

Wonseok Jung¹ · Young-Hwa Jeong¹ · Moon-Ghu Park² · Chang-Kyo Lee³ · Jeongwook Seo^{1*}

¹Namseoul University · ²Sejong University · ³GERI

E-mail : dnjstjr93@gmail.com / yhjeong@nsu.ac.kr / mgpark@sejong.ac.kr / ecg999@geri.re.kr / jwseo@nsu.ac.kr

요 약

본 논문에서 기상 데이터를 사용하여 태양광 에너지를 예측하기 위해 기계학습 모델인 SVM(Support Vector Machine)과 ANN(Artificial Neural Network)의 성능을 비교한다. 장·단파 복사선 평균, 강수량, 온도 등 15가지 종류의 기상 데이터를 사용하여 두 모델을 생성하고, 실험을 통해 최적의 SVM의 RBF(Radial Basis Function) 파라미터와 ANN의 은닉층과 노드 개수, 정규화 파라미터를 도출하였다. SVM과 ANN 모델의 성능을 비교하기 위한 지표로서 MAPE(Mean Absolute Percentage Error)와 MAE(Mean Absolute Error)를 사용하였다. 실험 결과 SVM 모델은 MAPE=21.11, MAE=2281417.65의 성능을 달성하였고 ANN은 MAPE=19.54, MAE=2155345.10776의 성능을 달성하였다.

ABSTRACT

In this paper, we compare the performances of SVM (Support Vector Machine) and ANN (Artificial Neural Network) machine learning models for predicting solar energy by using meteorological data. Two machine learning models were built by using fifteen kinds of weather data such as long and short wave radiation average, precipitation and temperature. Then the RBF (Radial Basis Function) parameters in the SVM model and the number of hidden layers/nodes and the regularization parameter in the ANN model were found by experimental studies. MAPE (Mean Absolute Percentage Error) and MAE (Mean Absolute Error) were considered as metrics for evaluating the performances of the SVM and ANN models. Simulation results showed that the SVM model achieved the performances of MAPE=21.11 and MAE=2281417.65, and the ANN model did the performances of MAPE=19.54 and MAE=2155345.10776.

키워드

Artificial Neural Network, Support Vector Machine, Prediction, Weather Data, Solar Energy

1. 서 론

태양광 발전 시스템은 태양광으로부터 에너지를 생산하여 부하에 공급하는 발전기술로 부하에 에너지를 안정적으로 공급하는 것이 가장 중요하다. 그러나 기상 조건에 따른 복합적인 요소가 미치는 영향으로 인해 에너지 생산이 불안정하기 때문에

안정적으로 부하를 운용하기 위해서는 에너지 생산량에 대한 정확한 예측이 필요하다 [1-2].

본 논문에서는 15가지 종류의 기상 데이터를 사용하여 보다 정확하게 태양광 에너지를 예측할 수 있도록 기계학습 모델인 SVM(Support Vector Machine)과 ANN(Artificial Neural Network)을 생성하고 실험을 통해 각 모델에 필요한 최적의 파라미터를 도출한다. 최종적으로 생성된 두 기계 모델에 대해 MAPE(Mean Absolute Percentage Error)와 MAE(Mean Absolute Error)를

* corresponding author

성능지표로 성능을 비교하여 태양광 에너지 예측에 적합한 모델을 선정한다.

II. 기상 데이터셋

태양광 에너지 예측을 위해 사용되는 데이터는 예측 모델 및 분석 대회 플랫폼 Kaggle의 2013년부터 2014년까지 개최된 태양 에너지 예측 경연 대회의 데이터이다. 1994년부터 2007년까지 총 14년 동안의 미국 오클라호마 주 Chickasha 지역의 기상 데이터로, 하루 단위 15개 기상 요소로 구성되고 데이터 수는 약 75,000개이다. 15가지 기상 요소의 자세한 설명은 표 1과 같다 [3].

표 1. 기상 요소 설명

Variable	Description	Units
apcp_sfc	3-Hour accumulated precipitation at the surface	kg m-2
dlwrf_sfc	Downward long-wave radiative flux average at the surface	Wm- 2
dswrf_sfc	Downward short-wave radiative flux average at the surface	Wm- 2
pres_msl	Air pressure at mean sea level	Pa
pwat_eatm	Precipitable Water over the entire depth of the atmosphere	kg m -2
spfh_2m	Specific Humidity at 2 m above ground	kg kg -1
tcdc_eatm	Total cloud cover over the entire depth of the atmosphere	%
tcopc_eatm	Total column-integrated condensate over the entire atmosphere	kg m -2
tmax_2m	Maximum Temperature over the past 3 hours at 2 m above the ground	K
tmin_2m	Minimum Temperature over the past 3 hours at 2 m above the ground	K
ttmp_2m	Current temperature at 2 m above the ground	K
ttmp_sfc	Temperature of the surface	K
ulwrf_sfc	Upward long-wave radiation at the surface	W m-2
ulwrf_tatm	Upward long-wave radiation at the top of the atmosphere	W m-2
uswrf_sfc	Upward short-wave radiation at the surface	W m-2
total	the total daily incoming solar energy	Jm- 2

III. 태양광 에너지 예측용 기계학습 모델

SVR 모델을 설계하기 위해 기상 데이터셋에서 훈련 및 테스트 데이터를 분할한 다음 15가지 종류의 기상 데이터 각각에 최소-최대 스케일링을 적용하여 표준화하였다. SVR 모델은 클래스 간의 마진을 최대화하는 초평면(Hyperplane)을 찾아냄으로써 클래스를 분류하는 방법이다. 제안한 모델은 자유 파라미터 γ 를 0.01로 설정하고 정규화 파라미터 C를 10^3 으로 설정하여 실험한다[4].

ANN 모델은 SVR 모델과 달리 데이터를 분할한 다음 스케일링을 적용하지 않는다. ANN 모델은

입력이 들어왔을 때 가중치를 이용하여 계산된 결과를 출력으로 내보내는 구조이며 훈련은 각 레이어의 가중치를 결정하는 과정이다. 제안한 모델은 은닉층을 16개의 노드와 10개의 레이어로 구성하고, 정규화 파라미터 α 를 10^{-5} 로 설정하여 큰 크기의 가중치를 페널티로 적용하여 과대적합을 방지한다 [5].

성능 비교를 위한 평가지표로 MAPE와 MAE를 사용한다. MAPE는 데이터에 극단이 없을 때 가장 잘 작동하는 척도이며 예측 오류에서 가장 일반적이다.

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (1)$$

두 개의 연속 변수의 차이를 나타내는 척도이며 각 점과 $Y=X$ 사이의 평균 수평 거리이다.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2)$$

여기서 y 는 실제 데이터 값이며 \hat{y} 은 예측값을 나타낸다.

IV. 실험 결과

태양광 에너지 예측을 위해 기상 데이터셋을 80%의 훈련 데이터, 20%의 테스트 데이터로 분할하고 설정한 파라미터를 사용하여 기계학습 모델을 생성 후 실험한 결과 표 2와 같이 SVR은 21.11의 MAPE, 2.28×10^6 의 MAE 성능을 확인하였고, ANN은 19.54의 MAPE, 2.15×10^6 의 MAE 성능을 확인하였다. 또한 그림 1과 그림 2는 테스트 데이터와 예측 결과 데이터의 비교 그림이다.

표 2. 성능평가 지표를 통한 기계학습 모델 비교

Model	MAPE	MAE
SVR	21.11	2281417.645
ANN	19.54	2155345.107

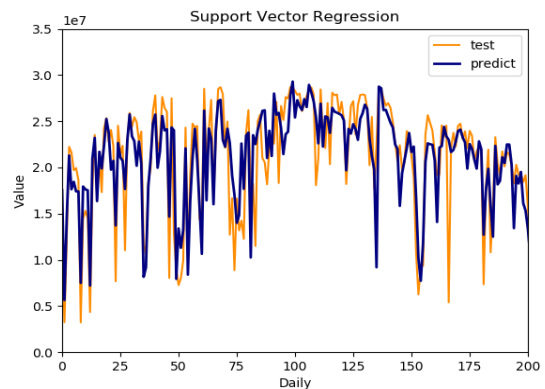


그림 1. 테스트 데이터와 SVR을 이용한 예측 결과

데이터의 비교

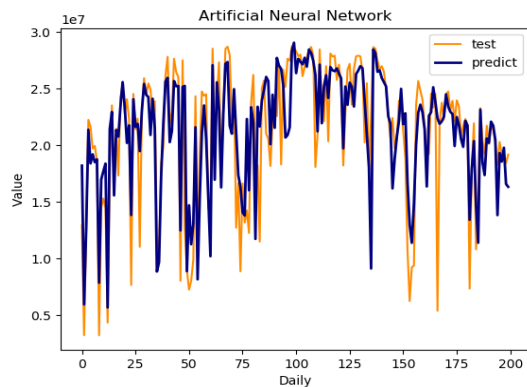


그림 2. 테스트 데이터와 ANN을 이용한 예측 결과 데이터의 비교

V. 결 론

본 논문에서는 기상 데이터셋을 이용한 기계학습 모델 비교를 통해 태양광 에너지를 예측하였다. 사용된 기계학습 모델은 SVR 모델과 ANN이며 실험을 통해 파라미터를 설정하고 비교하였다. γ 를 0.01, C를 10^3 으로 설정한 SVR 모델은 21.11의 MAPE와 2.28×10^6 의 MAE 성능을 확인하였고, 은닉층을 16개의 노드와 10개의 레이어로 구성하고 α 를 10^{-5} 로 설정한 ANN 모델은 19.54의 MAPE와 2.15×10^6 의 MAE 성능을 확인하였다.

Acknowledgement

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea (No. 20171210201140).

References

- [1] G. Y. Chun, Y. Han, S. Y. Song and J. H. Park, "Power estimation for stand-alone PV system Using forecast information," in *Proceeding of the Korean Institute of Electrical Engineers*, pp. 228-231, Apr. 2014.
- [2] K. H. Kim, S. W. Jo, R. J. Park and K. B. Song, "The Effect of Small Photovoltaic Power Generations and Illumination Intensity in Short-Term Electric Load Forecasting," in *Proceeding of the Korean Institute of Electrical Engineers*, Busan, pp. 461-462, Jul. 2017.

- [3] Kaggle. AMS 2013-2014 Solar Energy Prediction Contest [Internet]. Available : <https://www.kaggle.com/c/ams-2014-solar-energy-prediction-contest/>.
- [4] A. Müller and S. Guido, *Introduction to Machine Learning with Python*, O'Reilly Media, pp. 125-137, 2017.
- [5] Scikit-learn. MLPRegressor [Internet]. Available: http://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html/.