

순환신경망 기반 태양광 발전량 예측 시스템

정설령*, 박경욱**, 고진광*, 이성근*

*순천대학교, **전남대학교

411098@scnu.ac.kr, xelfiria@nate.com, kjg@scnu.ac.kr, sklee@scnu.ac.kr

Photovoltaic Prediction System based on Recurrent Neural Network

Seolryung Jung*, Kyoungwook Park**, Jingwang Koh*, Sungkeun Lee*

*Suncheon National University, **Chonnam National University

요 약

화석연료의 빈번한 사용으로 인한 지구온난화 문제가 심각해지면서 화석연료를 대체할 수 있는 신재생 에너지가 떠오르고 있다. 그중에서도 에너지원이 청정하고 무제한으로 사용할 수 있다는 장점을 가진 태양광 발전소가 주목을 받고 있다. 하지만 기후에 따라 영향을 많이 받는 특징 때문에 안정적인 전력 생산을 위해서는 태양광 발전량 예측이 매우 중요해지고 있다. 본 논문에서는 시계열 데이터에 특화된 순환신경망 기법인 RNN과 LSTM 모델을 이용하여 태양광 발전량을 예측하고 각 모델의 하이퍼 파라미터를 다르게 주어 비교 분석하였다. 그 결과 LSTM 모델이 RNN 모델보다 높은 예측력을 보였고, 손실 값이 0.1보다 낮은 높은 정확도를 보였다.

1. 서론

지금까지 석탄, 석유, 천연가스과 같은 화석연료는 인류가 이용하고 있는 가장 중요한 에너지 자원이 되었다. 하지만 화석연료의 사용으로 인한 환경오염이 세계적으로 문제가 되고 있다. 대표적으로 지구온난화를 일으키는 온실가스 배출의 주요 요인으로 화석연료가 꼽힌다. 따라서 오늘날에는 화석연료를 대체할 수 있는 새로운 에너지 자원으로 신재생 에너지가 대두되고 있다. 태양, 지열, 풍력, 수력 등 자연을 이용한 신재생 에너지는 화석연료에 비하여 오염물질의 배출이 적고 지속 가능한 에너지로 주목받고 있다. 그중에서도 태양의 빛을 이용한 태양광 에너지가 이슈되고 있다.

태양광 발전은 태양빛을 전기에너지로 변환시키는 발전 기술로서 태양빛과 패널 사이의 광전효과를 활용한 전기 생산방식이다[1]. 태양광 발전은 에너지원이 청정하고 무제한으로 사용할 수 있다는 장점이 있다[2]. 또한 20년 이상으로 장기적인 운용이 가능하며, 사람의 개입을 최소화 할 수 있는 무인화 시스템이 가능하다[3]. 하지만 설치 장소가 한정적이고 초기 설치비용이 높다는 단점이 있다. 또한 전력 생산량이 날씨에 의존적이며, 특히 일사량에 따른 변

동이 심하다. 일반 가정용·사업용 태양광 발전소 모두 태양광 에너지를 직접 활용하는 경우를 제외하고 소득을 낼 수 있는 방법은 전력 거래소에게 일정량의 에너지를 판매하는 수익구조라고 볼 수 있다. 따라서 전력 거래소에게 정보제공을 하기 전에 발전량의 예측이 선행되어야 효율을 극대화할 수 있다.

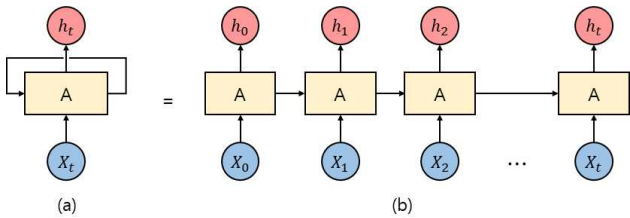
본 논문에서는 2019년부터 2021년까지의 시간별 데이터를 이용하여 태양광 생산량을 예측하는 순환신경망 시스템을 구축하려고 한다.

2. 관련연구

2.1 순환신경망

순환신경망(Recurrent Neural Network, RNN)[4]은 은닉 계층 안에 하나 이상의 순환 계층을 갖는 신경망 구조로 그림 2와 같은 구조를 갖는다. (a)는 기본적인 RNN 구조이고, (b)는 (a)를 펼쳐놓은 형태이다. x_t 는 입력 데이터, h_t 는 은닉 데이터를 나타낸다. 은닉층의 출력값을 출력층으로 전달해 주는 것과 동시에 은닉층 출력값이 다시 순환되어 은닉층에 입력값으로 들어간다. 이러한 순환신경망은 시계열 데이터를 처리하는 데 있어 강점을 갖는다. 하지

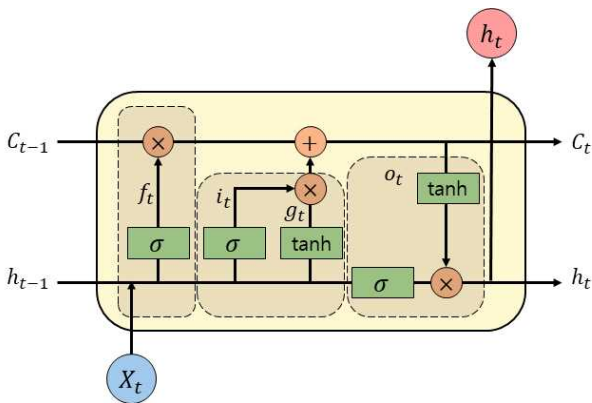
만 이전 시점에서 이어받은 데이터와 사용 지점의 거리가 멀 경우 과도한 역전파(back propagation)로 인해 기울기(gradient)가 소실(vanishing) 또는 발산(exploding)하는 문제가 발생한다[5]. 따라서 단기 기억은 유지되지만 장기기억력은 떨어지는 장기 의존성 문제(problem of Long-Term Dependency)가 나타난다. 이를 해결하기 위해 LSTM 모델이 등장했다.



(그림 1) RNN 구조.

2.2 LSTM(Long Short-Term Memory)

LSTM은 RNN의 은닉층에 메모리 셀과 3가지 게이트를 추가한 신경망이다. 그림 3은 LSTM의 구조이며 LSTM 셀 내부에서 동작하는 각 게이트들을 나타낸다. f_t 는 망각 게이트(forget gate)로 과거의 불필요한 기억은 지우고 필요한 기억만 남기는 비율을 조정한다. i_t 는 입력 게이트(input gate)로 새로운 기억을 추가하는 비율을 조정하며, o_t 는 출력 게이트(output gate)로 메모리 셀 내부를 출력에 반영하는 비율을 조정한다. 이와 같은 작동원리로 게이트들을 적절히 이용하여 장단기 기억을 활용할 수 있다.



(그림 2) LSTM 모델 구성.

3. 순환신경망 기반 태양광 발전량 예측 시스템

3.1 데이터 수집

예측에 사용된 데이터는 2019년부터 2021년까지 전라남도 영암지역의 태양광발전소 데이터를 이용했다. 일사량이 없는 시간대를 제외하고 평균적으로 6시~18시까지 매시간마다 관측된 평균 데이터를 사용

했다. 태양광 데이터에는 환경센서로 수집된 경사 일사량, 수평 일사량, 모듈 온도, 대기 온도와 패널 전력량이 포함된다. 경사 일사량은 태양광 패널과 수평인 환경 센서로 관측한 일사량을 말하며, 수평 일사량은 일반적으로 관측된 일사량이다. 모듈 온도는 패널의 온도로서 온도가 높아질수록 생산량이 저하되는 특징이 있다. 패널 전력값은 본 논문에서 예측하고자 하는 태양광 전력량을 나타낸다.

3.2 태양광 발전량 예측 시스템 구현 및 결과 분석

시계열 데이터에 특화된 순환신경망 기법 중 RNN 모델과 LSTM 모델을 이용하여 예측을 진행하였다. 먼저 2019~2020년도까지의 데이터는 훈련데이터, 2021년도 이후의 데이터는 테스트 데이터로 나누는 뒤 MinMaxScaler를 이용한 정규화를 통해 각 변수값마다 상이한 데이터를 같은 단위(0~1 사이)로 통일시켰다. 이후 훈련을 진행하는데, 이때 배치 사이즈(batch size), 시간 간격(time step), 훈련 횟수(epoch), 학습률(learning rate), 은닉층 수(number of hidden layer) 등을 조정하여 좀 더 정확도를 높일 수 있다. 본 논문에서 RNN 모델은 각각 은닉층을 2개, 3개로 구성하였고, 30의 타임 스텝을 주었다. LSTM은 각 모델마다 3개의 은닉 계층과 타임 스텝 50의 동일한 값으로 지정해 주고, 아래 표 1과 같이 최적의 모델을 도출하기 위해 나머지 하이퍼 파라미터(hyper parameter) 값들을 다르게 지정해 주었다. Units은 은닉층의 노드 수를 나타낸다. 학습은 노드 수를 200부터 800까지 진행한 결과 RNN 모델은 2가지 LSTM 모델은 4가지로 추려졌다.

<표 1> 하이퍼 파라미터 지정 단위.

Model	Units	Learning Rate	Batch Size	Epochs
RNN 1	200	1e-3	32	100
RNN 2	300	1e-3	32	200
LSTM 1	200	1e-3	20	200
LSTM 2	300	1e-3	20	100
LSTM 3	500	1e-3	64	50
LSTM 4	700	1e-3	64	50

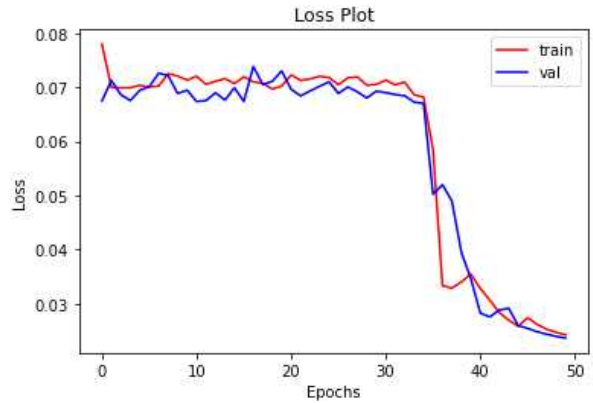
그림 3은 앞서 살펴보았던 LSTM 1번 모델과 4번 모델의 실제 태양광 전력 값과 예측값을 비교해 놓은 그래프다. x축은 2021년 이후의 날짜를 뜻하며 y

축은 태양광 패널 전력을 나타낸다. 빨간 선은 실제 전력량을 나타내고, 노란 선은 예측값을 나타낸다.

예측 결과 두 모델이 비슷한 정확도로 예측하고 있는 것을 볼 수 있다.

<표 2> 모델 성능 평가 지수(MSE, RMSE, R2).

Model	MSE	RMSE	R2
RNN 1	0.0465	0.2156	0.2140
RNN 2	0.0382	0.1954	0.3540
LSTM 1	0.0204	0.1428	0.6158
LSTM 2	0.0251	0.1587	0.5257
LSTM 3	0.0335	0.1830	0.4275
LSTM 4	0.0183	0.1353	0.6550



(그림 4) LSTM 4의 Loss 그래프.

4. 결론

태양광 발전은 전력 생산이 기후 변화에 크게 영향을 받는다는 문제점이 있어 발전량 예측의 중요성을 부가시킨다. 본 논문에서는 2019년부터 2021년까지의 시간별 데이터를 이용하여 태양광 생산량을 예측하는 순환신경망 시스템을 구축하였고, 그 결과 손실 값이 0.1보다 낮은 높은 예측 정확도를 보였다. 또한 순환신경망 기법인 RNN과 LSTM 모델을 비슷한 조건에서 비교해 보았을 때 LSTM이 더 좋은 성능을 보였다. 현재, 태양광 발전소에서 측정된 데이터만으로 예측을 진행하였다. 향후 연구에서는 날씨 변수와 같은 다양한 변수값을 추가하여 예측을 진행하고, RNN, LSTM뿐만 아니라 다양한 알고리즘을 적용하여 분석해 볼 예정이다.

ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성(Grand ICT 연구센터) 사업의 연구결과로 수행되었음” (IITP-2021-2020-0-01489)

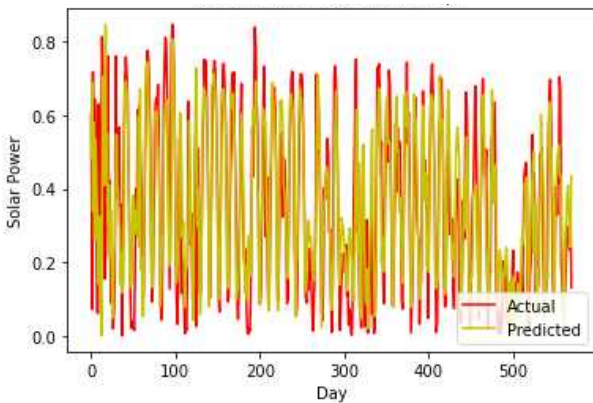
참고문헌

[1] S. Bracco, M. Brignone, F. Delfino, and R. Procopio, 2017. An energy management system for the savona campus smart polygeneration microgrid. IEEE Syst. J., vol. 11, no. 3, pp. 1799 - 1809, 2017.

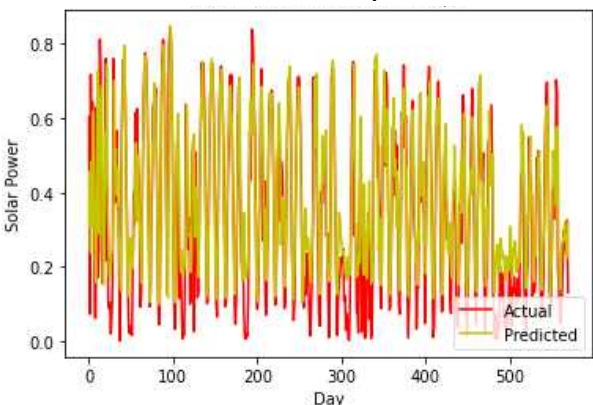
[2] S. Lee, “Let’s Replace fossil fuels with solar energy”, Proceedings of KIEE, vol. 49, no. 1, pp. 38-41, 2000.

[3] D. Kim, “[Special feature: Solar cell technology development trends] Domestic and overseas solar technology development and market status”,

Solar Power Prediction Graph - LSTM 1



Solar Power Prediction Graph - LSTM 4



(그림 3) 태양광 발전량 예측 그래프.

그림 4는 LSTM 4 모델의 Loss Plot이다. Loss값은 MSE로 설정하였다. 그 결과 Loss 값이 0.1 이하의 높은 정확도를 보여주었다.

따라서 LSTM 모델은 태양광 발전량 예측에 적합한 알고리즘이라고 할 수 있다.

Korean Solar Energy Society, vol. 4, no. 3, pp. 3-9, 2005.

[4] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back propagating errors," *nature*, vol. 323, no. 6088, pp. 533-536, 1986.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.