

기후 및 계절정보를 이용한 딥러닝 기반의 장기간 태양광 발전량 예측 기법

Deep Learning Based Prediction Method of Long-term Photovoltaic Power Generation Using Meteorological and Seasonal Information

이동훈(Donghun Lee)*, 김관호(Kwanho Kim)**

초 록

최근 온실가스의 증가로 인한 기후변화 대응의 필요성과 전력수요의 증가로 인해 태양광 발전량(PV) 예측의 중요성은 급격히 증가하고 있다. 특히, 태양광 발전량을 예측하는 것은 합리적인 전력 가격결정과 시스템 안정성 및 전력 생산 균형과 같은 문제를 효과적으로 해결하기 위해 전력생산 계획을 합리적으로 계획하는데 도움이 될 수 있다. 그러나 일사량, 운량, 온도 등과 같은 기후정보 및 계절 변화로 인한 태양광 발전량이 무작위적으로 변화하기 때문에 정확한 태양광 발전량을 예측하는 것은 도전적인 일이다. 따라서 본 논문에서는 딥러닝 모델을 통해 기후 및 계절정보를 이용하여 학습함으로써 장기간 태양광 발전량 예측 성능을 향상시킬 수 있는 기법을 제안한다. 본 연구에서는 대표적인 시계열 방법 중 하나인 계절형 ARIMA 모델과 하나의 은닉층으로 구성되어 있는 ANN 기반의 모델, 하나 이상의 은닉층으로 구성되어 있는 DNN 기반의 모델과의 비교를 통해 본 연구에서 제시한 모델의 성능을 평가한다. 실험데이터를 통한 실험 결과, 딥러닝 기반의 태양광 발전량 예측 기법이 가장 우수한 성능을 보였으며, 이는 본 연구에서 목표로 한 태양광 발전량 예측 성능 향상에 긍정적인 영향을 나타내었음을 보여준다.

ABSTRACT

Recently, since responding to meteorological changes depending on increasing greenhouse gas and electricity demand, the importance prediction of photovoltaic power (PV) is rapidly increasing. In particular, the prediction of PV power generation may help to determine a reasonable price of electricity, and solve the problem addressed such as a system stability and electricity production balance. However, since the dynamic changes of meteorological values such as solar radiation, cloudiness, and temperature, and seasonal changes, the accurate long-term PV power prediction is significantly challenging. Therefore, in this paper, we propose PV power prediction model based on deep learning that can be improved the PV

이 논문은 2017년도 정부(미래창조과학부) 한국연구재단이 후원하는 개인기초연구사업(No. NRF-2017R1D1A1B03035639) 지원을 받아 수행된 연구임.

* First Author, Department of Industrial and Management Engineering, Incheon National University (dhlee@inu.ac.kr)

** Corresponding Author, Department of Industrial and Management Engineering, Incheon National University (khokim@inu.ac.kr)

Received: 2018-11-09, Review completed: 2019-01-05, Accepted: 2019-01-29

power prediction performance by learning to use meteorological and seasonal information. We evaluate the performances using the proposed model compared to seasonal ARIMA (S-ARIMA) model, which is one of the typical time series methods, and ANN model, which is one hidden layer. As the experiment results using real-world dataset, the proposed model shows the best performance. It means that the proposed model shows positive impact on improving the PV power forecast performance.

키워드 : 태양광 발전량 예측, Deep Learning, Machine Learning, 시계열 분석, 계절형 ARIMA Model

Photovoltaic Power Prediction, Deep Learning, Machine Learning, Time Series Analysis, Seasonal ARIMA Model

1. 서 론

최근 온실가스의 증가로 인한 기후변화 대응의 필요성과 전력수요의 증가로 인해 태양광 발전량(Photovoltaic Power) 예측의 중요성은 급격히 증가하고 있다. 태양광 발전시스템은 노동력 없이 기후정보를 이용하여 전력을 자동 생산할 수 있는 장점이 있다. 또한, 태양광 발전 시스템을 통해 생산된 전력은 전력 공급자와 운영자에게 도움이 될 수 있다. 전력 공급자는 전력 공급과 수요 계획을 담당하고 있으며, 태양광 발전량 예측은 시스템 안정성 및 전력 생산 균형과 같은 문제를 효과적으로 해결하기 위한 전력생산 계획을 합리적으로 계획하는데 도움이 될 수 있다[12]. 전력 운영자는 전력 가격을 결정하고 시민에게 전력을 공급하는 역할을 담당한다. 태양광 발전량 예측은 사전에 전력의 가격을 결정하는데 도움이 될 수 있으며 합리적인 가격으로 시민에게 전력을 제공할 수 있다.

게다가 정확한 장기간 태양광 발전량을 예측하는 것은 태양광 발전 시스템을 운영하는 사업자의 수익을 최대화하기 위해 매우 중요하다. 대부분의 태양광 발전량의 가격은 하루 동안의

가격이 정해지는 시장과[23], 하루 중 시간에 따라 가격이 정해지는 시장이 존재한다[32]. 특히, 태양광 발전량의 가격은 태양광 발전량에 의한 많은 양의 전기가 생산되었을 때 가격이 상승한다. 그러므로 사업자의 수익은 얼마나 많은 전기가 생성될지 예측할 수 있는 능력에 직접적인 영향을 받는다. 따라서 정확한 장기간의 태양광 발전량의 예측은 태양광 발전량의 판매 유무를 결정함으로써 사업자의 수익을 최대화하는데 도움이 될 수 있다.

그러나 일사량, 운량, 온도 등과 같은 기상정보 변화로 인한 태양광 발전량의 변화로 인해 정확한 태양광 발전량을 예측하는데 어려움이 존재한다. 예를 들어, 봄과 여름에는 일사량이 비교적 높고 태양이 늦게 지기 때문에 태양광 발전량을 얻을 수 있는 시간과 태양광 발전 추정량이 많아진다. 반면에 가을과 겨울에는 일사량이 비교적 낮고 태양이 빨리 지기 때문에 태양광 발전 추정량이 적어진다. 또한 기상 조건들은 시간 및 지역 기후조건에 따라 역동적이고 무작위로 변화하며 이에 따라 태양광 발전량도 무작위로 변화한다. 게다가 대부분의 태양광 발전량 예측시스템은 일사량만을 고려하여 계산하도록 발전되었으나[6, 15, 18] 실제

국내에 설치된 태양광 발전소의 발전량의 측정량을 조사해 본 결과, 같은 지역이라 할지라도 동일한 발전량을 보이지 않았다[15]. 따라서 장기간 태양광을 예측하기 위해서는 일사량뿐만 아니라 다양한 기상정보 및 계절정보를 고려해야 하는 어려움이 존재한다.

이와 같은 문제를 해결하기 위해 태양광 발전량을 예측하기 위한 시계열 분석과 인공지능 기반의 기법을 통해 많은 연구가 진행되었다. 예를 들어, 시계열 분석 기법 중 하나인 Autoregressive Integrated Moving Average (ARIMA) 모델을 이용하여 태양광 발전량 예측 정확도를 평가하였다[24]. 또한 제한적인 환경에서 기상정보 데이터를 이용하여 Artificial Neural Network(ANN) 모델 기반의 단시간 태양광 발전량을 예측하였고 일사량과 온도정보 데이터를 이용하여 Deep Neural Network (DNN) 모델 기반의 단기 태양광 예측 기법에 대해 연구하였다[28]. 그리고 단기 예측뿐만 아니라 기상정보를 이용하여 DNN 모델 기반의 태양광 발전량 예측 기법에 대한 연구도 진행되었다[16]. 그러나 기존연구에서 태양광 발전

량은 계절에 따라 많은 차이가 있음에도 불구하고 계절적 정보를 고려하지 않고 일사량 정보만을 이용하거나 기상정보만을 이용하여 장기 태양광 발전량을 예측하거나 단기 태양광 발전량을 예측하는 한계가 존재하였다.

따라서 본 논문에서는 장기간 태양광 발전량 예측 성능을 향상시키기 위해 일사량뿐만 아니라 온도, 습도, 일사량 등의 기상정보와 계절정보를 이용하여 계절정보를 반영하지 못했던 문제점을 해결하고 딥러닝 모델 기반의 태양광 발전량 예측 기법을 제안한다. 실 데이터를 통한 실험결과, 본 연구에서 제안한 기후 및 계절정보를 이용한 딥러닝 기반의 기법이 가장 우수한 성능을 보였으며, 이는 본 연구에서 목표로 한 태양광 발전량 예측 성능 향상에 긍정적인 영향을 나타내었음을 보여준다.

본 논문의 구성은 다음과 같다. 제2장에서 관련연구에 대해 설명하고, 제3장에서는 문제정의와 제안한 모델에 대한 내용을 설명한다. 제4장에서는 제안한 태양광 발전량 모델에 대한 성능을 평가하고 마지막으로 제5장에서 결론을 기술한다.

〈Table 1〉 Overview of Previous Research Related to PV Power Prediction

Features considered	Analysis methods	References
Short-term prediction	ARIMA	Fernandez-Jimenez et al.[9], Pedro and Coimbra[24]
	SVM	da Silva Fonseca et al.[5], Shi et al.[25]
	ANN	Lee[19], Izgi et al.[11], Sulaiman et al.[26]
	DNN	Wang et al.[28], Ashraf and Chandra[1]
Long-term prediction	ARMAX, ARIMA	Li et al.[20], Li et al.[21]
	MAR	Li et al.[20]
	ANN	Ding et al.[7], Chen et al.[4]
	DNN	Kou et al.[16]

2. 관련 연구

기존연구는 태양광 발전량 단기간 예측과 장기간 예측 측면에서 <Table 1>과 같이 연구되었다. 이전 연구들에서 단기간에 대한 정의를 4시간 예측과[14] 하루 동안의 태양광 발전량을 예측으로 정의하였으나[27], 또 다른 이전 연구에서는 장기간을 하루 동안으로 정의하였다[31]. 따라서 본 연구에서는 이전 연구들을 고려하여 단기간을 하루 이전의 시간으로 정의하였고, 장기간을 하루 또는 하루 이상으로 정의하였다. 먼저 태양광 발전량 단기간 예측 연구에 대해 살펴보면, ARIMA, K-nearest neighbors(KNN), ANN 기반의 모델 등을 이용하여 태양광 발전량 예측 비교를 통해 평가하였고[24], 머신러닝 기법 중 Support Vector Machine(SVM)을 이용한 태양광 발전량 예측 기법을 제시하였다[5, 25]. 특히 기상정보를 분류하여 예측함으로써 성능을 향상시켰다[25]. 그리고 기상정보를 이용한 인공지능망 모델 중 하나인 ANN 기반의 모델 기반의 태양광 발전량 예측 기법에 대해 연구하였고[11, 26], 은닉층의 신경망을 층으로 쌓아 만든 DNN 기반의 모델 기반의 단기 예측

기법을 제시하였다[1, 28].

다음으로 장기간 예측 측면에서 보면, 기상정보를 이용하여 Exogenous Autoregressive Moving Average(ARMAX), ARIMA 모델 등 통계적 시계열 분석 기법을 통해 태양광 발전량 예측에 대해 연구하였다[20, 21]. 특히 Li et al.[20]의 연구에서는 시계열 분석 모델뿐만 아니라 추가적으로 Multivariate Adaptive Regression (MAR), KNN모델들과의 성능 비교를 통해 연구를 확장시킴으로써 장기 예측 분석 기법에 대해 연구하였다. 날씨 유형에 따라 분류된 과거 유형과 가장 유사한 기상 유형을 날씨 데이터를 분류 한 후 ANN 기반의 모델을 통해 예측하였고[7], ANN 기반의 모델 중 Radial Basis Function Network(RBFN)를 사용하여 태양광 발전량을 예측하였다[4]. 마지막으로 기상정보를 기반으로 DNN 기반의 모델을 이용하여 태양광 발전량 장기 예측에 대해 연구하였다[16]. 하지만 기존 연구들에서는 기상정보만을 이용하여 시계열 분석 기법과 인공지능망 모델을 활용하여 태양광 발전량을 예측하는데 초점을 두고 있다. 따라서 본 연구에서는 계절정보를 추가적으로 사용하여 계절적 특성을 반영하지

<Table 2> Input and Output Features and Description for Long-Term PV Power Prediction

	Category	Variables	Description
Input values	Meteorological data ($x_i, i = 1, \dots, n$)	Temperature(e)	e_i is temperature at the i -th hour
		Humidity(h)	h_i is humidity at the i -th hour
		Cloudiness(c)	c_i is cloudiness at the i -th hour
		Radiation(r)	r_i is radiation at the i -th hour
		Month of year(u)	u_i is month of year at the i -th hour
		Day of month(g)	g_i is day of month at the i -th hour
Output value	PV power	Predicted PV power(\hat{y})	\hat{y}_i is predicted PV power at the i -th hour
		Actual PV power(y)	y_i is actual PV power at the i -th hour

못했던 문제점을 해결하고 딥러닝 기반의 장기간 태양광 발전량 예측 기법을 제안한다.

3. 제안 기법

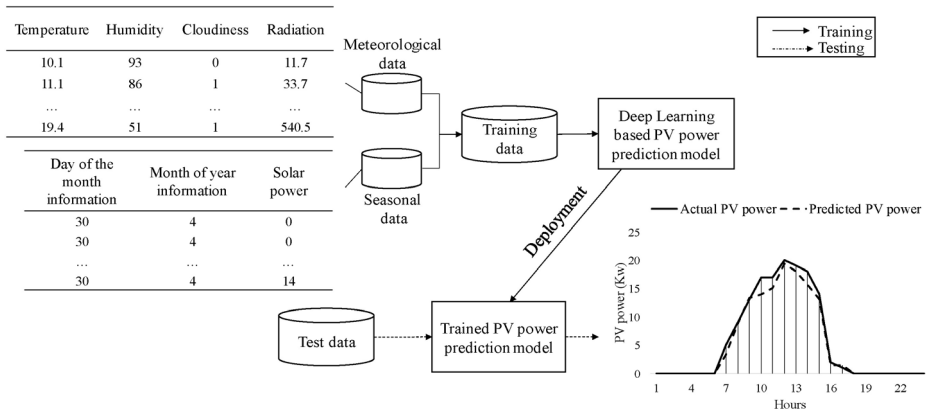
3.1 문제 정의

우리의 연구목표는 기후 및 계절정보를 이용하여 제안한 모델을 통해 장기간 태양광 발전량을 예측하는 것이다. <Table 2>는 장기간 태양광 발전량 예측을 위해 고려되어진 변수를 나타낸다. 먼저 제안한 모델의 학습에 사용되는 입력 값은 기후 및 계절정보가 있다. 기후정보로는 온도(e), 습도(h), 운량(c), 일사량(r)이 있고 계절정보는 월정보(u), 일정보(g)가 있다. 출력 값으로는 예측된 태양광 발전량(\hat{y})와 실제 태양광 발전량(y)이 있다. $i = 1, \dots, n$,는 시간의 정보를 나타낸다. $vec(x)_i$, $i = 1, \dots, n$,는 하루 중 i 번째 시간에서의 기후 및 계절정보로 이루어진 $\langle e_i, h_i, c_i, r_i, u_i, g_i \rangle$ 의 벡터를 나타낸다. \hat{y}_i , $i = 1, \dots, n$,는 학습된 모델에서 하루

중 i 번째 시간에서의 예측된 태양광 발전량을 나타내고 y_i , $i = 1, \dots, n$,는 하루 중 i 번째 시간에서의 실제 태양광 발전량을 나타낸다.

3.2 장기간 태양광 발전량 예측 모델 프레임 워크

<Figure 1>은 기후 및 계절정보를 이용하여 신경망 모델 학습 방법을 통한 장기간 태양광 발전량을 예측하기 위한 제안된 프레임워크를 나타낸다. 프레임워크는 학습 과정과 테스트 과정으로 구성된다. 프레임워크 모델은 입력 값, 활성화함수, 손실함수 및 최적화 기법을 기반으로 신경망 모델을 통해 학습한다. 신경망 모델은 입력층, 은닉층, 출력층으로 구성되어있다. 학습 과정에서 사용되는 입력 값 $vec(x)_i$ 는 각각의 층에 있는 연결 가중치와 활성화함수와의 계산을 통해 비선형 값으로 변환된다. 우리는 파라미터를 학습하기 위해 Backpropagation 기법을 [10] 적용하였고 예측된 태양광 발전량과 실제 태양광 발전량과의 오차를 최소화하는 연결 가중치 값을 찾기 위해 반복적으로 학습한다.



<Figure 1> Proposed Framework for Long-Term PV Power Prediction

테스트 과정에서는 학습 과정에서 학습된 연결 가중치를 사용한다. 연결 가중치는 학습 과정에서 사용하지 않는 입력 값 $vec(x)_i$ 를 이용하여 학습된 모델을 기반으로 예측된 태양광 발전량과 실제 태양광 발전량과의 차이를 계산하여 태양광 발전량 예측의 성능을 평가한다.

3.3 계절형 ARIMA 기반의 태양광 발전량 예측 모델

ARIMA 모델은 Auto-Regressive(AR)과 Moving Average(MA) 모델을 결합한 시계열 분석 기법이다. 시계열들은 시간에 따라 변하는 확률구조, 특히 평균과 공분산의 행태가 시간에 따라 달라지는 특징을 갖는다. 계절형 ARIMA(S-ARIMA) 모델은 ARIMA 모델과 주기의 자료들 간의 관계를 동시에 반영한 모형으로 계절성을 보이는 시계열 분석에 널리 사용된다. 따라서 태양광 발전량은 기후 및 계절정보에 따라 영향을 많이 받기 때문에, 기후 및 계절정보를 고려하여 S-ARIMA 기반의 모델을 통해 태양광 발전량을 예측하고 평가한다.

ARIMA 모델에서 예측 방정식에 나타나는 차분된 계열의 시차는 AR항이라고 부르며 예측 오차의 시차를 MA항이라고 부른다. $ARIMA(p, d, q)$ 모델은 $AR(p)$ 모델과 $MA(q)$ 모델과 합친 $ARMA(p, q)$ 모델에 잔차를 제거하고 식 (1)과 같이 표현한다. 잔차는 $\Delta X_t = X_t - X_{t-1} = (1-B)X_t$ 로 정의한다.

$$(1 - \phi_1 B - \dots \phi_p B^p)(1 - B)^d X_t \quad (1) \\ = (1 - \phi_1 B - \dots \phi_q B^q) a_t$$

여기에서 정수형 인덱스 t 를 가진 어떤 시계열, X_t 를 유한한 개수의 자신의 과거의 값들에 상수를 곱한 것과 정규분포를 따르는 백색소음(white noise), a_t 의 합으로 나타낸다. B 는 자연연산자(lag operator)이고, p 와 q 는 유한한 지연 차수를 나타낸다[30].

따라서 S-ARIMA(p, d, q)(P, D, Q) $_z$ 은 식 (2)와 같이 나타낸다.

$$\phi(B)\Phi(B^z)(1-B)^d(1-B^z)^D X_t \quad (2) \\ = \Theta(B)\theta(B^z)a_t$$

여기서 z 는 주기의 인덱스로 12개월을 주기로 나타낸다고 가정할 경우, 12가 된다. 학습에 사용되는 데이터를 기반으로 기후 및 계절정보에 대한 입력 값과 장기간의 태양광 발전량에 대한 출력 값과의 시계열적인 관계에 대해 근본적인 영향을 미치는 요소를 실용적인 접근법을 이용하여 식 (1), 식 (2) 계산을 통해 S-ARIMA 모델의 파라미터 p, d, q, P, D, Q, z 를 최적의 값으로 결정한다[8, 29]. 따라서 최적의 파라미터를 통해 학습에 사용하지 않은 데이터를 기반으로 태양광 발전량을 예측하고, 실제 태양광 발전량과의 차이를 통해 성능을 평가한다.

3.4 ANN 기반의 태양광 발전량 예측 모델

ANN 기반의 모델은 예측, 인지 등 많은 분야에 적용되어 높은 성능을 보였다[2, 13]. 이 모델은 입력층, 은닉층, 출력층으로 이루어져 있는 인공신경망 모델이며, 우리는 은닉층이 하나인 ANN 기반의 모델을 제안한다. 이 모델은 입력 값 $vec(x)_i$ 를 이용하여 가중치, 활성화함수, 손실

함수 및 최적화 기법을 통해 실제 태양광 발전량과의 차이를 최소화하기 위해 반복적으로 학습한다.

먼저 입력 값 $vec(x)_i$ 에 대해 노드가 입력 값을 수신하면 입력 값의 가중 합계를 간단히 계산한다. 여기서, 각 입력 값에 대한 가중치는 초기에는 0과 1 사이의 무작위 값으로 설정되고, 학습 단계 동안 업데이트 된다. 특히, 노드 q 에 대한 값은 net_q 로 표시되며 식 (3)을 사용하여 계산된다.

$$net_k = \sum_{j=1}^k w_{k,j} o_j \quad (3)$$

여기에서 $w_{k,j}$ 는 노드 j 에서 k 로의 연결에 해당하는 가중치이고 o_j 는 이전 계층에서 노드 j 가 생성한 출력 값이며, j 는 노드 k 에 연결된 노드수이다. 다음으로 노드 k 에 대한 값을 사용하여, f_k 로 표시되는 출력 값을 식 (4)를 통해 계산한다.

$$f_k = \begin{cases} v(net_k), & \text{if } net_k > 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

여기에서 $v(\cdot)$ 는 활성화 함수를 나타낸다. 출력층에서 노드의 최종 출력 값인 예측된 태양광 발전량 f_k , $k = 1, \dots, n$ 로 하루 중 k 시간의 태양광 발전량이 출력된다. 우리는 학습 파라미터를 최적화하기 위해 사용되는 활성화 함수와 최적화기법을 Rectified Linear Unit(ReLU)와 Adaptive Moment Estimation(Adam)을 사용하였고[17] 학습 과정에서 예측된 태양광 발전량과 실제 태양광 발전량과의 차이를 계산하기 위해 손실함수는 식 (5)와 같이 Mean Square Error(MSE)함수를 사용하였다.

$$L(x) = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (5)$$

최종적으로 식 (5)를 통해 예측한 태양광 발전량 \hat{y}_i , $i = 1, \dots, n$ 과 실제 태양광 발전량 y_i , $i = 1, \dots, n$ 과의 오차를 최소화하는 연결 가중치 값을 찾기 위해 Backpropagation기법을 기반으로 출력층에서 입력층으로 직선방향으로 각 연결 가중치를 변경함으로써 반복적으로 학습한다. 따라서 학습된 연결 가중치 값은 학습에 사용하지 않은 데이터를 기반으로 기후 및 계절 정보를 이용하여 태양광 발전량을 예측하고, 실제 태양광 발전량과의 차이를 통해 성능을 평가한다.

3.5 DNN 기반의 태양광 발전량 예측 모델

DNN 기반의 모델은 은닉층의 수가 다양한 ANN 기반의 모델을 기반으로 과거의 기후 및 계절정보를 학습하여 태양광 발전량을 예측하는 인공신경망 모델이다. DNN 기반의 모델은 은닉층의 수가 증가함에 따라 매개변수의 수가 증가하기 때문에 학습하는데 어려움이 있었으나 Backpropagation 알고리즘으로 문제점을 해결함으로써 ANN 기반의 모델보다 더 높은 성능을 보인다. 우리는 반복적인 실험을 통해 은닉층의 개수를 7개로 설정하였다. 이 모델은 ANN 기반의 모델과 동일한 활성화함수, 손실 함수, 최적화 기법을 사용하여 실제 태양광 발전량과의 오차를 최소화하기 위해 반복적으로 학습한다. 따라서 ANN 기반의 모델과 마찬가지로 학습에 사용하지 않은 데이터를 기반으로 기후 및 계절 정보를 이용하여 태양광 발전량을 예측하고, 성능을 평가한다.

4. 실험 및 평가

4.1 실험 환경

본 연구에서 제안한 딥러닝 기반의 모델을 학습하기 위해 사용한 기상정보 데이터는 한국 기상청에서 제공하는 데이터를 수집하였다. <Table 3>은 수집된 데이터의 예를 보여준다. 계절적 정보 중에서 월의 정보는 각각의 월의 값을 나타낸다. 마찬가지로 일의 정보는 각각의 일의 값을 나타낸다. 만약 8월 1일인 경우 월의 값은 8이고 일의 값은 1일이 된다. 습도의 범위는 0에서 100이며 높을수록 습도가 높다는 것을 나타낸다. 운량의 범위는 0에서 10이며 높을수록 운량이 많아 날씨가 흐리다는 것을 나타낸다. 입력 값과 출력 값의 n 은 하루 중 태양광 발전량이 측정되는 시간 06:00에서 19:00까지 측정되므로 하루의 태양광 발전량을 예측하기 위해 14로 설정하였다.

제안한 모델을 학습하기 위한 데이터는 2013년 1월부터 2015년 2월까지의 데이터를 사용하였고 학습 시 유효성을 검증하기 위한 데이터는 2015년 3월부터 2015년 8월까지의 데이터를 사용하였다. 모델의 성능을 검증하기 위한 테스트 데이터는 2016년 3월, 4월, 8월, 10월의 데이터를 사용하였다. S-ARIMA 기반의 모델은 실용적인 접근법을 이용하여 4월(2, 0, 13)(1, 0, 2)₄, 8월(2, 0, 13)(1, 0, 2)₄, 10월(2, 0, 2)(1, 0, 2)₄, 3월(2, 0, 13)(1, 0, 2)₄으로 각각 설정하였다 [8,29]. 또한, 제안한 DNN 기반의 모델에서 은닉층의 수에 따른 모델의 성능을 평가하기 위해 은닉층의 수가 4개인 C_DNN을 모델을 추가적으로 설정하여 실험하였다.

본 연구에서 제안하는 예측 모델들이 얼마나 실제 태양광 발전량을 잘 추정하는지 정량화하기 위해 주어진 입력 값에 대한 예측된 태양광 발전량과 실제 발전량과의 차이를 MAE 기반으로 식 (6)을 통해 계산한다. 여기에서 m 은

<Table 3> Obtained Input and Output Values for Learning Proposed Models (August 11, 2015)

Hour	Temperature (℃)	Humidity (%)	Cloudiness Index	Radiation (W/m ²)	Month of year (-)	Day of month (-)	Actual PV Power (Kw)
06	23.4	93	1	0	M	D	0
07	24.3	90	1	93.1	M	D	2
08	26	85	1	249.7	M	D	5
09	27.5	78	1	346.1	M	D	6
10	29.7	68	1	408.9	M	D	9
11	30.8	65	1	447.1	M	D	11
12	33.6	52	2	677.7	M	D	18
13	33	55	3	656.5	M	D	20
14	33.1	53	3	663.9	M	D	18
15	31.4	55	4	592.5	M	D	17
16	32.1	53	3	423.5	M	D	11
17	30.9	62	3	323.4	M	D	8

시간별 예측의 횟수를 나타낸다. MAE 로 설정한 이유는 실제 태양광 발전량보다 예측한 태양광 발전량이 적을 경우, 음수의 값이 되므로 u 에 따라 모델의 성능이 우수한 것처럼 나타날 수 있기 때문에 절대 값을 사용하여 성능을 평가한다.

$$MAE = \frac{1}{m} \frac{\sum_{i=1}^m |\hat{y}_i - y_i|}{\sum_{i=1}^m y_i}, (i = 1, \dots, m) \quad (6)$$

4.2 모델 학습과정

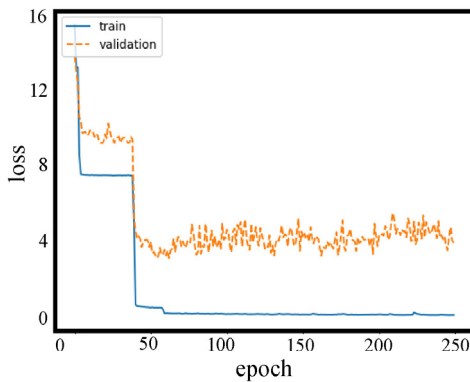
<Figure 2>는 ANN과 DNN 기반의 모델의 학습과정을 나타낸다. 먼저 ANN과 DNN 기반의 모델을 학습 할 때 배치 크기를 128로 설정하였다. 손실 함수는 배치 크기에 대해 손실 값을 계산하므로 더 많은 예측 태양광 발전량과 실제 태양광 발전량과의 손실 값을 동시에 고려 할 수 있고 학습 속도가 빠르다. 또한 성능은 전체 태양광 발전량 오차를 손실 함수를 통해 계산하는 것과 유사한 성능을 보인다[22]. ANN과

DNN 기반의 모델의 학습 반복 횟수는 250으로 설정하였다. 학습 반복 횟수가 많으면 많은 과적합(Overfitting)으로 인해 모델의 성능이 저하될 수도 있고, 학습 반복 횟수가 적으면 소적합(Underfitting)으로 인해 모델의 성능이 저하가 발생하여 성능이 저하될 수 있다. 따라서 우리는 반복적인 실험을 통해 학습 반복 횟수를 설정하였다.

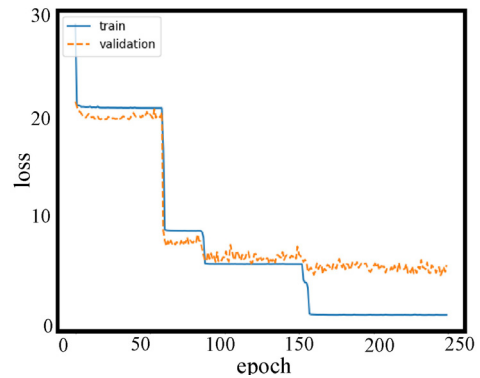
<Figure 2>와 같이 학습 반복 횟수에 따라 제안한 모델들이 학습하면서 학습 및 검증 손실 값의 변화를 나타낸다. 그래프에서 나타나 있는 바와 같이, 학습 반복 횟수가 증가할수록 학습 및 검증 손실 값이 감소하였으며, 이는 주어진 모델들이 태양광 발전량을 예측하기 위한 방법을 찾기 위해 성공적으로 학습되었음을 나타낸다.

4.3 실험평가 및 분석

<Figure 3>은 본 연구에서 고려 된 예측 모델의 전반적인 성능을 비교한 것을 나타낸다.



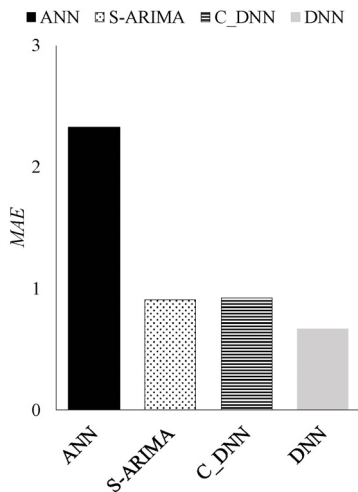
(a) ANN based model



(b) DNN based model

<Figure 2> Changes of the Loss Values According to the Number of Epochs During the Training Phase

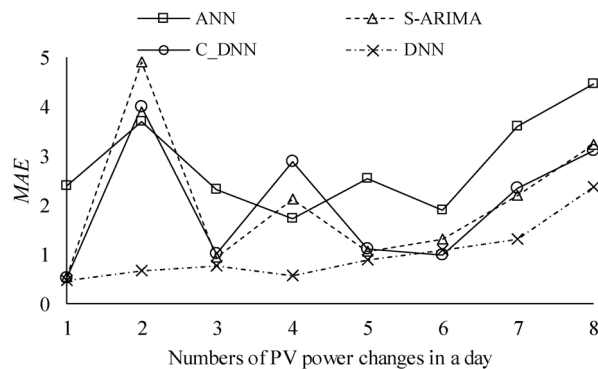
시계열 패턴을 고려하지 않고 현재 상태에만 초점을 맞춰 학습된 ANN 기반의 모델은 기존 시계열 예측 방법인 S-ARIMA 기반의 모델에 비해 효과가 적었다. 반면에 ANN 기반의 모델을 개선한 DNN 기반의 모델은 ANN 기반의 모델, S-ARIMA 기반의 모델, C_DNN 모델에 대해 각각 71.2%, 26.4%, 27.6%의 *MAE*를 감소시켰다.



〈Figure 3〉 Performance Comparisons Using Testing Dataset in Terms of *MAE*

특히, <Figure 4>는 태양광 발전량 변화에 대해 단기적인 관점에서 성능을 평가한 것을 나타낸다. 여기에서는 태양광 예측 난이도에 따라 성능을 비교하기 위해 매일의 태양광 발전량 출력 변화의 수를 계산한다.

ANN 기반의 모델은 태양광 발전량 변화 수와 관계없이 가장 좋지 않은 성능을 보인다. 반면에, DNN 기반의 모델은 모든 경우에 대해 기존 모델에 비해 상당히 낮은 오류를 보인다. 특히, 난이도가 높을수록 예측 모델의 오류가 커지는 경향이 있다. S-ARIMA 기반의 모델과 ANN 기반의 모델은 태양광 발전량의 출력 변화가 높은 날에는 3% 이상의 오류가 많은 결과를 나타내고, DNN 기반의 모델은 3% 미만으로 오류가 적다. 따라서 실험을 통한 결과는 하루에 단기간의 시간별 패턴을 포착하여 개선된 학습과정을 통해 학습한 DNN 기반의 모델은 태양광 발전량의 변화가 거의 없을 때뿐만 아니라 태양광 발전량의 변화가 많을 경우에도 높은 예측을 성능을 보인다는 것을 확인하였다. 그러나 태양광 발전량의 변화에 연속적인 관측을 통한 통계적 측면에서 예측한 S-ARIMA 기반의 모델은 태양광 발전량 변화가 많은 경우



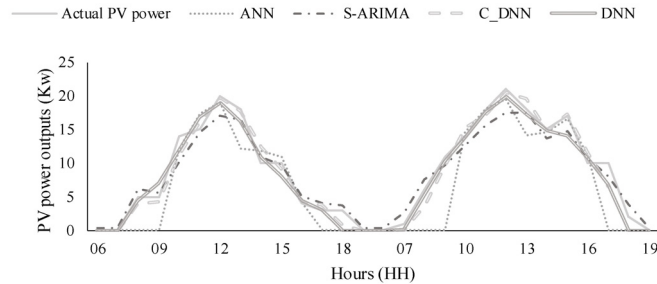
〈Figure 4〉 The *MAE* Results According to the Changes of PV Power in a Day

태양광 발전량을 정확히 예측하지 못한다는 것을 확인할 수 있었다.

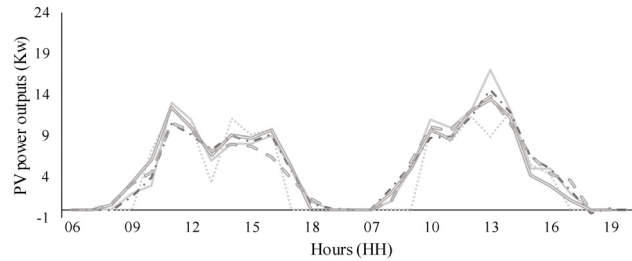
<Figure 5>는 연속 2일 동안의 태양광 발전량 장기적인 변화에 따라 3가지 관점에서 실제 태양광 발전량과 예측한 태양광 발전량을 보여 준다. 첫 번째, 태양광 발전량이 일정할 때 피크 시간대의 높은 태양광 발전량은 날씨가 맑고 일사량이 높음을 나타낸다. 기후와 계절적 정보가 2일 동안 유사하기 때문에 태양광 발전량

은 장기적으로 변화가 거의 없으며, 제안된 모델들은 비슷한 성능을 갖지만 피크시간대의 S-ARIMA 기반의 모델이 가장 오류가 많이 발생하고 실제 값보다 낮을 것으로 예측한다.

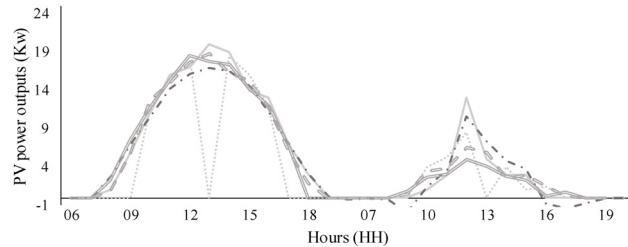
반면에 <Figure 5>에 있는 (b)와 (c)와 같이 제안된 모델들의 성능차이는 태양광 발전량이 장기적으로 급격히 변할 때 나타난다. 특히 피크 시간대에 태양광 발전량의 변화가 많이 발생한다. 태양광 발전량이 급격히 떨어지면 DNN



(a) When PV power outputs are similar(August, 17, 18, 2016)



(b) When PV power outputs rapidly rise(October, 17, 18, 2016)



(c) When PV power outputs rapidly drop(March, 3, 4, 2016)

<Figure 5> Visualization of the Predicted PV Power Outputs for Consecutive Two Days

기반의 모델은 피크 시간대에 가장 낮은 성능을 보인다. S-ARIMA 기반의 모델은 실제 태양광 발전량과 가장 유사하게 예측한다. 반대로 태양광 발전량이 급격히 상승하면 ANN 기반의 모델은 가장 낮은 성능을 보인다. DNN 기반의 모델은 태양광 발전량이 급격히 떨어지는 경우보다 나은 성능을 보이며, 가장 우수한 성능을 보인다.

이 실험을 통해 태양광 발전량에 따라 제안된 모델의 성능에 차이가 있음을 확인하였다. 실제로 기후 변화는 예상하기 힘들만큼 급격히 변화하기 때문에 태양광 발전량의 변화가 크다. 또한 태양광 발전량은 전기의 수요와 가격의 측면에서 낮았을 때보다 높았을 때 태양광 발전량을 정확히 예측하는 것이 중요하다. 비록 기후 정보에 계절적 정보가 반영이 될지라도 4계절을 가지고 있는 한국에서는 계절에 따라 기후가 심하게 변하는 특성이 있으며, <Figure 5>의 (b)와 (c)와 같이, 태양광 발전량이 급격히 변화하는 경향을 보인다. 따라서 한국에서 태양광 발전량을 잘 예측하기 위해서는 태양광 발전과 기상변화의 관계뿐만 계절적 요소를 고려해야한다.

S-ARIMA 기반의 모델의 경우 기상변화와 태양광 발전량의 변화를 통계적으로 분석한 결과를 기반으로 태양광 발전량을 예측하였으나, 태양광 발전량이 급격히 높아졌을 때 성능이 좋지 않았음을 확인하였다. 반면에, DNN 기반의 모델은 과거의 기상 및 계절 정보에 대한 태양광 발전량 정보를 반복적으로 학습함으로써 장기간 태양광 발전량을 예측하는데 높은 성능을 보였다. 따라서 장기간 태양광 발전량 예측 성능을 향상시키기 위해서는 기후정보뿐만 아니라 계절정보를 고려하는 것이 필수적이라는 것을 확인하였다.

5. 결 론

본 연구에서는 기후 및 계절정보를 이용한 딥러닝 기반의 장기간 태양광 발전량 예측 분석 기법을 제안하였다. 본 연구에서 제안한 모델들의 성능을 평가하기 위해 2가지 방법을 제시하였다. 첫 번째, 4계절 4주 동안의 제안한 모델에 따른 예측 태양광 발전량과 실제 태양광 발전량과의 오차를 통해 성능을 평가하였다. 두 번째, 태양광 발전량이 피크시간에 높게 출력되었을 때와 낮게 출력되었을 때의 예측을 통해 성능을 평가하였다. 실험을 통해 본 연구에서 제안한 모델들의 성능을 평가한 결과 태양광 발전량을 잘 예측하기 위해서는 태양광 발전량의 영향을 끼치는 과거 기후 정보뿐만 아니라 계절 정보를 이용하여 반복적으로 학습이 필요한 모델이 필요하며, 학습 모델 중 은닉층의 개수가 두 개 또는 두 개 이상인 딥러닝 기반의 모델을 통한 태양광 발전량 예측이 필수적이라는 것을 확인하였다.

본 연구의 결과는 전력의 공급과 수요를 계획하는 공급자와 전력의 가격을 결정하는 운영자에게 보다 정확하게 태양광 발전량을 예측하여 정확한 생산계획과 가격결정을 가능하게 할 것이다. 향후 연구에서는 태양광 발전량 예측의 성능을 향상시키기 위해 딥러닝 모델 중 은닉층에 있는 노드들의 방향을 가지고 연결되어 순환구조를 이루는 Recurrent Neural Network(RNN) 모델을 적용하여 태양광 발전량을 예측하거나 RNN모델에서 의 데이터 학습 시 학습능력이 크게 저하되어 이 문제를 극복하기 위해 고안된 Long Short-term Memory(LSTM) 모델을 적용한 태양광 발전량 예측 분석 연구가 필요하다.

References

- [1] Ashraf, I. and Chandra, A., "Artificial Neural Network Based Models for Forecasting Electricity Generation of Grid Connected Solar PV Power Plant," *International Journal of Global Energy Issues*, Vol. 21, No. 1-2, pp. 119-130, 2004.
- [2] Bae, J. K., Lee, S. Y., and Seo, H. J., "Artificial Intelligence Techniques for Predicting Online Peer-to-Peer(P2P) Loan Default," *The Journal of Society for e-Business Studies*, Vol. 23, No. 3, pp. 207-224, 2018.
- [3] Cha, W. C., Park, J., Cho, U., and Kim, J. C., "Design of Generation Efficiency Fuzzy Prediction Model Using Solar Power Element Data," *The Transactions of The Korean Institute of Electrical Engineers*, Vol. 63, No. 10, pp. 1423-1427, 2014.
- [4] Chen, C., Duan, S., Cai, T., and Liu, B., "Online 24-h Solar Power Forecasting Based on Weather Type Classification Using Artificial Neural Network," *Solar Energy*, Vol. 85, No. 11, pp. 2856-2870, 2011.
- [5] da Silva Fonseca Jr, J. G., Oozeki, T., Takashima, T., Koshimizu, G., Uchida, Y., and Ogimoto, K., "Use of Support Vector Regression and Numerically Predicted Cloudiness to Forecast Power Output of a Photovoltaic Power Plant in Kitakyushu, Japan," *Progress in Photovoltaics: Research and Applications*, Vol. 20, No. 7, pp. 874-882, 2012.
- [6] Detyniecki, M., Marsala, C., Krishnan, A., and Siegel, M., "Weather-based Solar Energy Prediction," *WCCI 2012 IEEE International conference on Fuzzy Systems*, pp. 1-7, 2012.
- [7] Ding, M., Wang, L., and Bi, R., "An ANN-based Approach for Forecasting The Power Output of Photovoltaic System," *Proceeding of Environmental Sciences*, Vol. 11, No. 1, pp. 1308-1315, 2011.
- [8] Edward, G., Box, P., and Jenkins, G. M., "Time Series Analysis: Forecasting and Control," *The Journal of Technometrics*, Vol. 37, pp. 238-242, 1995.
- [9] Fernandez-Jimenez, L. A., Muñoz-Jimenez, A., Falces, A., Mendoza-Villena, M., Garcia-Garrido, E., Lara-Santillan, P. M., Zorzano-Alba, E., and Zorzano-Santamaria, P. J., "Short-term Power Forecasting System for Photovoltaic Plants," *Renewable Energy*, Vol. 44, pp. 311-317, 2012.
- [10] Hagan, M. T. and Menhaj, M. B., "Training Feedforward Networks with the Marquardt Algorithm," *Proceeding of IEEE Transactions on Neural Networks*, Vol. 5, No. 6, pp. 989-993, 1994.
- [11] Izgi, E., Öztöpal, A., Yerli, B., Kaymak, M. K., and Şahin, A. D., "Short-mid-term Solar Power Prediction by Using Artificial Neural Networks," *Solar Energy*, Vol. 86, No. 2, pp. 725-733, 2012.

- [12] Inman, R. H., Pedro, H. T., and Coimbra, C. F., "Solar Forecasting Methods for Renewable Energy Integration," *Progress in Energy and Combustion Science*, Vol. 39, No. 6, pp. 535-576, 2013.
- [13] Jung, H. I., Park, I. S., and Ahn, H., "Identifying the Key Success Factors of Massively Multiplayer Online Role Playing Game Design using Artificial Neural Networks," *The Journal of Society for e-Business Studies*, Vol. 17, No. 1, pp. 23-38, 2012.
- [14] Kardakos, E. G., Alexiadis, M. C., Vagropoulos, S. I., Simoglou, C. K., Biskas, P. N., and Bakirtzis, A. G., "Application of Time Series and Artificial Neural Network Models in Short-term Forecasting of PV Power Generation," In *Proceedings of the 48th International Universities Power Engineering Conference*, pp. 1-6, 2013.
- [15] Kim, D. H. and Kim, J. O., "The Solar Power with Weather and Generator Scheduling," *KIEE Summer Conference*, p. 131, 2008.
- [16] Kou, J., Liu, J., Li, Q., Fang, W., Chen, Z., Liu, L., and Guan, T., "Photovoltaic Power Forecasting Based on Artificial Neural Network and Meteorological Data," *Proceeding of IEEE Region 10 Conference*, pp. 1-4, 2013.
- [17] Kingma, D. P. and Ba, J., "Adam: A method for stochastic optimization," *arXiv pre-print arXiv:1412.6980*, 2014.
- [18] Lee, I. R., Bae, I. S., Jung, C. H., Kim, J. O., and Shim, H., "Photovoltaic Generation System Output Forecasting Using Irradiance Probability Distribution Function," *KIEE Summer Conference*, pp. 548-550, 2004.
- [19] Lee, H., "The Development of The Predict Model for Solar Power Generation Based on Current Temperature Data in Restricted Circumstances," *Journal of Digital Contents Society*, Vol. 17, No. 3, pp. 157-164, 2016.
- [20] Li, Y., He, Y., Su, Y., and Shu, L., "Forecasting The Daily Power Output of a Grid-connected Photovoltaic System Based on Multivariate Adaptive Regression Splines," *Applied Energy*, Vol. 180, No. 15, pp. 392-401, 2016.
- [21] Li, Y., Su, Y., and Shu, L., "An ARMAX Model for Forecasting the Power Output of a Grid Connected Photovoltaic System," *Renewable Energy*, Vol. 66, No. 1, pp. 78-8, 2014.
- [22] Li, M., Zhang, T., Chen, Y., and Smola, A. J., "Efficient Mini-batch Training for Stochastic Optimization," In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 661-670, 2014.
- [23] Liu, G., Xu, Y., and Tomsovic, K., "Bidding Strategy for Microgrid in Day-ahead Market Based on Hybrid Stochastic/ Robust Optimization," *IEEE Transactions on Smart Grid*, Vol. 7, No. 1, pp. 227-237, 2016.

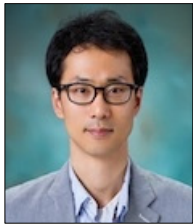
- [24] Pedro, H. T. and Coimbra, C. F., "Assessment of Forecasting Techniques for Solar Power Production with no Exogenous Inputs," *Solar Energy*, Vol. 86, No. 7, pp. 2017-2028, 2012.
- [25] Shi, J., Lee, W. J., Liu, Y., Yang, Y., and Wang, P., "Forecasting Power Output of Photovoltaic Systems Based on Weather Classification and Support Vector Machines," *IEEE Transactions on Industry Applications*, Vol. 48, No. 3, pp. 1064-1069, 2012.
- [26] Sulaiman, S. I., Rahman, T. A., and Musirin, I., "Partial Evolutionary ANN for Output Prediction of a Grid-Connected Photovoltaic System," *International Journal of Computer and Electrical Engineering*, Vol. 1, No. 1, pp. 40-45, 2009.
- [27] Tao, C., Shanxu, D., and Changsong, C., "Forecasting Power Output for Grid-connected Photovoltaic Power System without Using Solar Radiation Measurement," In *Proceedings of the International Symposium on Power Electronics for Distributed Generation Systems*, pp. 773-777, 2010.
- [28] Wang, S., Zhang, N., Zhao, Y., and Zhan, J., "Photovoltaic System Power Forecasting Based on Combined Grey Model and BP Neural Network," In *Proceedings of International Conference on Electrical and Control Engineering*, pp. 4623-4626, 2011.
- [29] Yule, G. U., "Why do We Sometimes Get Nonsense-Correlations Between Time-Series?: A Study in Sampling and the Nature of Time-series," *Journal of the Royal Statistical Society*, Vol. 89, No. 1, pp. 1-63, 1926.
- [30] Yu, L., Zhou, L., Tan, L., Jiang, H., Wang, Y., Wei, S., and Nie, S., "Application of a New Hybrid Model with Seasonal Autoregressive Integrated Moving Average (ARIMA) and Nonlinear Autoregressive Neural Network (NARNN) in Forecasting Incidence Cases of HFMD in Shenzhen, China," *PloS one*, Vol. 9, No. 6, 2014.
- [31] Yona, A., Senjyu, T., Funabashi, T., and Kim, C. H., "Determination Method of Insolation Prediction with Fuzzy and Applying Neural Network for Long-term Ahead PV Power Output Correction," *IEEE Transactions on Sustainable Energy*, Vol. 4, No. 2, pp. 527-533, 2013.
- [32] Zhou, Y., Wang, C., Wu, J., Wang, J., Cheng, M., and Li, G., "Optimal Scheduling of Aggregated Thermostatically Controlled Loads with Renewable Generation in the Intraday Electricity Market," *Applied Energy*, Vol. 188, pp. 456-465, 2017.

저 자 소 개



이동훈
2016년
2017년~현재
관심분야

(E-mail: dhlee@inu.ac.kr)
단국대학교 산업공학과 (학사)
인천대학교 산업경영공학과 석사과정
텍스트 마이닝, 통계적 기계학습



김관호
2006년
2012년
2013년
2014년~현재
관심분야

(E-mail: khokim@inu.ac.kr)
동국대학교 정보시스템전공 (학사)
서울대학교 산업공학과 (박사)
경희대학교 (연구박사)
인천대학교 산업경영공학과 교수
통계적 기계학습, 빅데이터 분석