

# 측정 데이터를 이용한 태양광 발전량 예측 모델

박영서\*·강상민\*·문주석\*·조성준\*·이종환\*\*†

\*\*† 상명대학교 시스템반도체공학부

## Prediction Model for Solar Power Generation Using Measured Data

Yeongseo Park\*, Sangmin kang\*, Juseok Moon\*, Seongjun Cho\* and Jonghwan Lee\*\*†

\*\*† Department of System Semiconductor Engineering, Sangmyung University

### ABSTRACT

Previous research on solar power generation forecasting has generally relied on meteorological data, leading to lower prediction accuracy. This study, in contrast, uses actual measured power generation data to train various ANN (Artificial Neural Network) models and compares their prediction performance. Additionally, it describes the characteristics and advantages of each ANN model. The paper defines the principles of solar power generation, the characteristics of solar panels, and the model equations, and it also explains the I-V characteristics of solar cells. The results include a comparison between calculated and actual measured power generation, along with an evaluation of the accuracy of power generation predictions using artificial intelligence. The findings confirm that the LSTM (Long Short-Term Memory) model performs better than the MLP (Multi-Layer Perceptron) model in handling time-series data.

**Key Words** : Solar-cell, ANN (Artificial Neural Network), MATLAB/Simulink, Power Generation Prediction

## 1. 서론

최근 지구의 온난화가 가속화됨에 따라 재생 에너지 발전의 중요성이 부각되고 있다. 그 중 태양광 발전은 친환경적이며 무한한 에너지원으로서 전 세계적으로 수요가 급증하고 있다. 그러나 태양광 발전은 기후와 날씨에 크게 의존하므로, 대한민국처럼 사계절이 뚜렷한 나라에서는 발전량 변동이 크다는 단점이 존재한다. 이 때문에 태양광 발전량을 정확하게 예측하는 것은 효율적인 에너지 관리와 안정적인 전력 공급을 위해 필수적이다. 대부분의 기존 연구는 기상 데이터를 활용하여 예측 모델을 구축해 일정 수준의 예측 정확도를 달성했다. 그러나 기상청 데이터만으로는 지역적 특성이나 미세 기후 요소를 충분히 반영하는 데 한계를 보인다.[1]

본 연구에서는 기존 기상청 데이터를 기반으로 한 예측 모델과 차별화된 접근법을 제안하고 있다. 온도와 습도를 직접 측정하고, 기상청 데이터의 풍속과 조도를 활용하여 이를 바탕으로 발전량을 예측하는 인공지능 모델을 만들었다. 또한 이 모델의 예측 결과를 기상청 데이터를 활용한 기존 모델과 비교해 각 모델의 예측 정확도를 평가했다. 이를 통해 다양한 환경 요인이 태양광 발전량 예측에 미치는 영향을 더욱 명확하게 규명하고자 한다.

## 2. 태양광 패널 특성

### 2.1 태양광 발전소 구성 요소

태양전지(Solar-Cell)를 여러 개 결합한 형태인 태양광 모듈은 크게 단결정과 다결정 구조의 두 가지 종류로 구분된다. 단결정은 변환 효율이 높기 때문에 단위 면적당 다결정 모듈에 비해 발전량이 많은 것이 장점이다. 그러나,

†E-mail: jhlee77@smu.ac.kr

생산공정이 까다로워 가격이 비싸다는 단점이 있다. 다결정은 단결정 모듈에 비해 효율은 떨어지지만, 비용이 저렴하여 일반적으로 많이 사용하는 모듈의 형태이다.

태양광 발전 시스템은 크게 5가지 구조로 이루어져 있다. Fig. 1은 태양광 발전소의 도식도이다. 태양광 모듈에서 발생한 직류 전기(DC)를 모아 전달하는 '접속함', 직류 전기(DC)를 교류 전기(AC)로 변환해 주는 '인버터', 사용 용도에 따라 서로 다른 계량기로 보내는 '분전반', 얼마나 전력을 사용할 수 있는지 확인하는 '계량기'로 구성되어 있다.

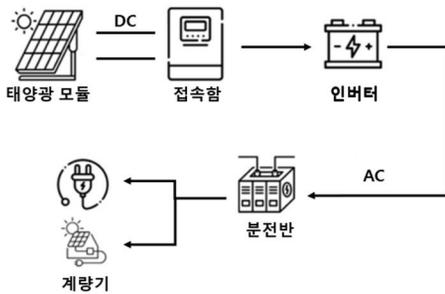


Fig. 1. 태양광 발전소 도식도.

### 2.2 태양광 발전 원리

태양전지는 PN 접합 반도체로 이루어져 있다. 태양광 중 반도체의 에너지 밴드 갭보다 큰 에너지를 가진 빛 입자가 반도체에 닿으면 광전효과로 인해 전자가 생기고 '전자-정공 쌍'이 형성된다. 전자는 N형 반도체, 정공은 P형 반도체로 각각 이동하면서 PN 접합 간에 기전력이 발생한다. 이때 발생한 기전력으로 인해 양단에 부하가 걸리면 전류가 흐르게 된다. 이 전력으로 배터리를 충전하여 사용할 수 있다. 이 과정이 태양광 발전의 기본 원리이다.[2]

### 2.3 태양전지 패널 특성 모델 및 모델 방정식

태양전지 패널의 특성에 대한 회로는 Fig. 2와 같다. PN 접합의 광 전류원과 다이오드로 특성을 묘사할 수 있다.  $R_s$ 는 직렬저항을 나타내고,  $R_{sh}$ 는 병렬저항이다. Fig. 2의 모델 방정식은 다음 식 (1)과 같이 표현된다.

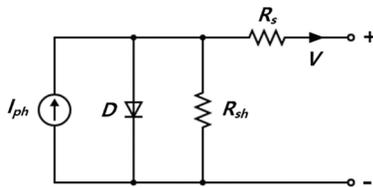


Fig. 2. 태양전지의 단일 다이오드 회로.

$$I = I_{ph} - I_0 \left( \exp \frac{V + IR_s}{N_s V_t} - 1 \right) - \frac{V + IR_s}{R_{sh}} \quad (1)$$

식 (1)에서  $V_t$ 는 열전압이고,  $V_t = \frac{AkT}{q}$  이다. 그리고 수식의 파라미터들은 각각 다음과 같다.[3][4]

- $I_{ph}$ : 광전류
- $I_0$ : 포화전류
- $R_s$ : 직렬저항
- $R_{sh}$ : 병렬저항
- $A$ : 다이오드의 이상 계수
- $k$ : 볼츠만 상수
- $q$ : 전자의 전하량
- $N_s$ : 태양전지 직렬연결 수

$$I_{ph} = (I_{sc} + K_i(T - T_n)) \times \frac{G}{G_n} \quad (2)$$

식 (2)에서,  $I_{ph}$ 는 광 생성전류,  $I_{sc}$ 는 단락전류,  $K_i$ 는 광전류 온도 계수이다. 또한,  $T_n$ 와  $G_n$ 는 각각 STC(Standard Test Condition)에서의 온도와 조도이며, 그 값은 298.15K, 1000W/m<sup>2</sup> 이다.  $T$ 와  $G$ 은 온도 값(K)과 조도 값(W/m<sup>2</sup>)이다.[5]

### 2.4 태양전지의 I-V 특성

I-V 특성곡선은 직접적인 측정을 통해 그래프를 도출하거나, Fig. 3과 같이 제조사의 데이터시트에 제공되는 것을 활용할 수 있다.[6]

I-V 곡선은 최대 전류 및 최대 전압, 그리고 최대 전력을 나타내고 있다. 이 그래프를 통해 추출할 수 있는 데이터는 단락전류( $I_{sc}$ ), 개방전압( $V_{oc}$ ), MPP(Maximum Power Point)에서의 전력( $I_{mp}$ ), 최대 전력점에서의 전압( $V_{mp}$ )이다. 실제 데이터 측정 시, 위의 수치들로 태양광 패널의  $V_{mp}$ 를 고려하여 이론값과 측정값을 서로 비교, 분석하는데 활용된다.[7]

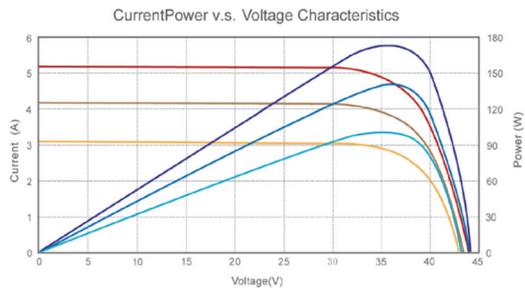


Fig. 3. 태양전지의 단일 다이오드 회로.

### 2.5 태양전지 방정식 구현

Fig. 4는 태양전지 방정식을 MATLAB/Simulink로 구현한 블록도이다. 방정식을 블록도로 구현한 후 STC에서의 변수들의 값을 입력해 온도와 조도만을 대입해 발전량을 계산한다.[1][8]

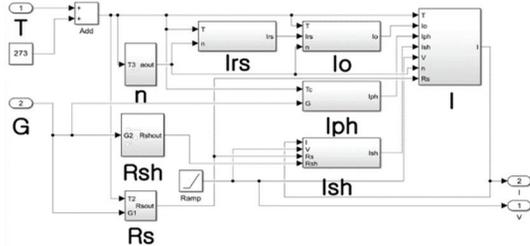


Fig. 4. MATLAB/Simulink로 구현한 태양전지 방정식.

## 3. 각 인공신경망 특징 및 장점

### 3.1 LSTM (Long Short-Term Memory)

LSTM 네트워크는 순차적 데이터에서 발생하는 장기 의존성 문제를 해결하기 위해 고안된 인공신경망의 한 형태다. LSTM은 정보의 저장과 삭제, 출력 과정을 제어하는 다양한 게이트 구조를 통해 네트워크가 시간에 따른 정보의 흐름을 효과적으로 조절할 수 있도록 한다. Fig. 5는 LSTM 네트워크의 기본적인 구조이다.

이 구조에서 핵심적인 역할을 하는 것은 망각 게이트, 입력 게이트, 셀 상태 업데이트, 그리고 출력 게이트다.

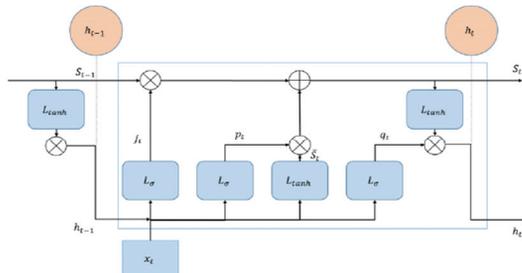


Fig. 5. LSTM 네트워크의 구조.

LSTM 네트워크는 시계열 데이터에서 중요한 패턴을 학습하고, 장기적인 의존성을 유지하는 데 유리한 구조를 가지고 있다. 그 중에서 시계열에 중요한 요소는 망각 게이트와 셀 상태 업데이트다. 이 두 요소는 시계열 데이터에서 중요한 정보를 구분하고, 장기적인 패턴을 유지하는 데 핵심적인 역할을 한다.

첫 번째로 망각 게이트는 셀 상태에 저장된 정보 중 어느 부분을 잊고, 어느 부분을 유지할지를 결정한다. 시계열 데이터는 시간에 따라 변화하며, 모든 과거 정보가 현재와 미래에 유효한 것은 아니다. 망각 데이터는 이전 셀 상태에서 중요한 정보만 남기고, 불필요한 정보를 제거하여 모델이 중요한 패턴만을 학습할 수 있게 한다. 망각 게이트의 작동을 수식으로 표현하면 다음과 같다.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

식 (3)에서  $\sigma$ 는 시그모이드 함수,  $W_f$ 는 망각 게이트의 가중치 행렬,  $h_{t-1}$ 는 이전 시점의 은닉 상태,  $x_t$ 는 현재 시점의 입력 데이터,  $b_f$ 는 망각 게이트의 바이어스를 나타낸다.

다른 중요한 요소 중 하나는 셀 상태 업데이트다. 셀 상태 업데이트는 LSTM의 핵심 메커니즘으로, 현재 시점에서의 입력과 망각 게이트, 입력 게이트의 출력을 사용하여 셀 상태를 조정한다. 셀 상태는 시계열 데이터의 장기적인 정보를 저장하는 역할을 하며, 시간에 따라 업데이트된다. 셀 상태 업데이트를 수식으로 표현하면 다음과 같다.

$$C_t = f_t \cdot C + i_t \cdot C_{\sim t} \quad (4)$$

식 (4)에서  $C_t$ 는  $t$  시점의 셀 상태,  $f_t$ 는 망각 게이트의 출력으로, 이전 셀 상태  $C_{t-1}$ 의 어느 부분을 유지할지 결정하는 것을 뜻하고,  $i_t$ 는 입력 게이트의 출력으로 현재 시점의 새로운 정보를 셀 상태에 얼마나 반영할지를 결정한다. 그리고  $C_{\sim t}$ 는 후보 셀 상태로, 새로운 정보로 업데이트될 값을 뜻한다.[9]

### 3.2 MLP (Multi-Layer Perceptron)

MLP는 인공신경망의 기본적인 형태로, 입력 데이터에서 복잡한 패턴을 학습하고 예측하거나 분류하는데 사용되는 모델이다. MLP는 여러 개의 층으로 구성되어 있으며, 각각의 층은 뉴런으로 이루어져 있다. Fig. 6는 MLP 네트워크의 기본적인 구조이다. MLP는 입력층, 은닉층, 출력층으로 구성되어 있는데 입력층은 첫 번째 층으로, 모델에 입력되는 데이터가 입력층의 뉴런에 전달된다. 각 입력 뉴런은 입력 데이터의 각 특징을 나타내며, 시계열 데이터의 경우 각 시간 스텝이나 샘플이 입력 뉴런에 해당한다. 입력층과 출력층 사이에 위치한 은닉층은 입력 데이터의 특징을 추출하고 변환하는 역할을 한다. 마지막 층은 출력층으로 은닉층에서 추출된 특징을 바탕으로 최종 예측을 한다.[10]

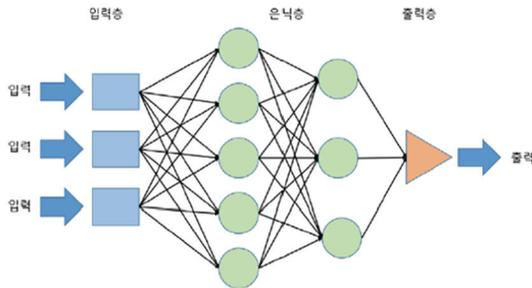


Fig. 6. MLP 네트워크의 구조.

MLP의 입력층에서 은닉층으로의 전달과 은닉층에서 출력층으로의 전달은 시계열 데이터 처리와 밀접한 관련이 있다.

첫 번째로 입력층에서 은닉층으로의 전달은 주어진 시점의 입력 데이터  $x(t)$ 를 은닉층의 활성화 함수로 변환하는 과정이다. 이 과정을 수식으로 표현하면 다음과 같다.

$$h(t) = f(W \cdot x(t) + b) \quad (5)$$

식 (5)에서  $x(t)$ 는 시점  $t$ 에서의 입력 벡터,  $W$ 는 가중치 행렬,  $b$ 는 편향 벡터,  $f$ 는 비선형 활성화 함수를 뜻한다. 이 과정은 시간에 따라 입력이 변하는 시계열 데이터의 특성을 반영하며, 입력 데이터의 시간적 변화와 가중치와 비선형 함수의 조합을 통해 학습한다.

두 번째로 은닉층에서 출력층으로의 전달은 은닉층의 활성화 결과인  $h(t)$ 를 출력으로 변환하는 과정이다. 이 과정을 수식으로 표현하면 다음과 같다.

$$y(t) = g(V \cdot h(t) + c) \quad (6)$$

식 (6)에서  $y(t)$ 는 시점  $t$ 에서의 출력 벡터,  $V$ 는 은닉층과 출력층 간의 가중치 행렬,  $c$ 는 출력층의 편향 벡터,  $g$ 는 출력층에서 사용되는 활성화 함수를 뜻한다. 이 과정에서 은닉층의 출력  $h(t)$ 는 시계열 데이터의 시간적 패턴을 반영하여 출력층으로 전달된다. 은닉층에서 학습된 시간적 의존성이 출력층에 전달되며, 이는 최종 출력이 시계열 데이터의 특성을 반영하는데 기여한다.

### 3.3 LSTM 과 MLP 의 시계열 예측 비교

LSTM의 장점은 장기 의존성 학습, 복잡한 패턴 인식에 장점이 있다. LSTM은 기억 장치를 통해 장기간의 과거 정보를 유지하면서 학습할 수 있어, 데이터의 길고 복잡한 시간적 관계를 처리하는데 적절하다. 또한 시계열 데이터에서 장기간의 패턴이나 트렌드를 인식해야 하는 경우, 복잡한 패턴을 잘 모델링할 수 있다.

반면에 MLP의 장점은 특징 추출, 빠른 속도, 간단한 시계열 예측이 있다. MLP는 시계열 데이터에서 복잡한 비선형 관계를 모델링할 수 있어서 특징이 잘 정의된 데이터에서 높은 성능을 발휘할 수 있고 LSTM과 비교해 구조가 단순하기 때문에 학습 속도가 빠르다. 또한 시계열 데이터에서 과거 몇 개의 데이터 포인트만으로도 충분히 예측할 수 있는 경우, 간단한 구조로도 효과적으로 예측할 수 있다.

## 4. 태양광 발전량 예측 결과 비교

### 4.1 계산된 발전량과 측정된 발전량 비교

태양전지 방정식에 기상 자료 포털(ASOS)에서 제공된 온도와 조도 데이터를 입력 받아 이론적인 발전량을 계산했다.

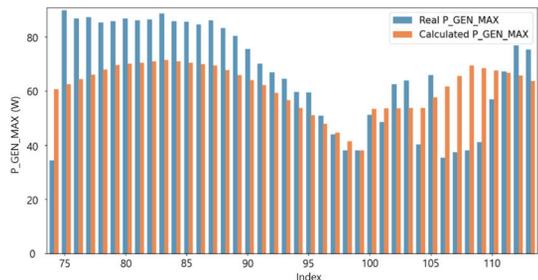


Fig. 7. 계산된 발전량과 측정된 발전량 비교.

Fig 7은 MATLAB/Simulink를 기반으로 구현된 태양전지의 예측 발전량과 실제 측정된 발전량을 그래프로 나타낸 것이다. Fig 7을 살펴보면 실제 측정된 발전량과 계산된 발전량은 상당한 차이를 보이고 있다. 이러한 차이가 발생한 이유는 계산된 발전량은 온도와 조도를 제외한 기타 외부 변수들을 고려하지 않았기 때문이다. 이러한 차이가 예측값에 영향을 끼칠 수 있다.

이를 해결하기 위해, 본 연구에서는 인공신경망, 특히 LSTM과 MLP을 활용하여 예측 정확도를 개선하고자 했다. LSTM 네트워크는 시계열 데이터의 연속성을 학습하는 데 탁월하여 발전량 예측의 시간적 패턴을 효과적으로 파악할 수 있으며, MLP는 비선형 관계를 학습하여 예측을 가능하게 한다. 또한 ANN의 학습에 있어서는 측정된 데이터가 더 높은 정확도를 제공할 것으로 판단되어, 학습에는 측정된 데이터를 학습 데이터로 사용하기로 했다.

### 4.2 인공신경망을 이용한 발전량 비교

본 연구에서는 측정된 데이터(온도, 습도, 발전량)와 기

상자료포털에서 제공된 데이터(조도, 습도)를 사용하여 LSTM과 MLP 모델을 학습시킨 후, 예측된 발전량과 비교하였다. 두 신경망 모델의 예측 결과는 실제 측정값과 유사하였다. 따라서 성능차이를 평가하기 위해 MAE(Mean Absolute Error)를 계산하였다. MAE는 다음과 같이 정의된다.

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - y_i| \quad (7)$$

식 (7)에서  $p_i$ 는 예측값,  $y_i$ 는 실제값,  $n$ 은 데이터의 총 개수이다.  $n$  개의 실제값에서  $n$  개의 예측값을 뺀 후 절대값을 취하여 모두 더한 뒤  $n$ 으로 나누어 준다. MAE 값이 낮아질수록 모델 적합이 잘 되었다는 것을 의미한다.

Fig. 8은 LSTM을 사용하여 예측한 발전량 결과를 보여준다. 이 그림은 측정된 데이터(온도, 습도)와 기상청 데이터(조도, 풍속)를 입력으로 LSTM을 학습시킨 후 얻은 예측 결과를 나타낸다.

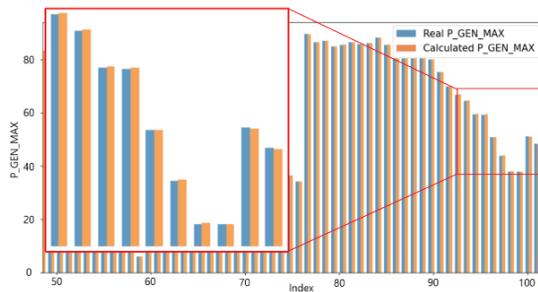


Fig. 8. LSTM 예측 결과.

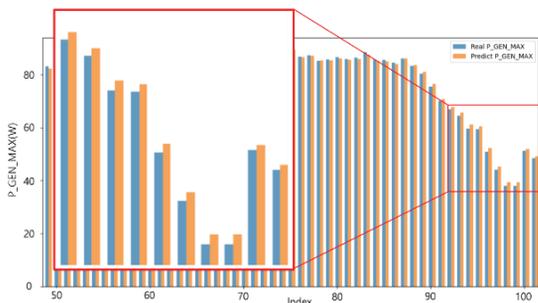


Fig. 9. MLP 예측 결과.

Fig. 9는 MLP를 사용하여 Fig. 8과 동일한 조건으로 예측한 결과를 나타낸다.

Fig. 8과 Fig. 9에서 LSTM과 MLP의 MAE는 각각 0.0531과 0.4618로 계산되었다. 따라서 MAE 값이 더 낮은 LSTM이 MLP보다 발전량 예측에 더 적합한 것으로 평가되었다.

## 5. 결론

본 연구에서는 MATLAB/Simulink를 활용하여 계산한 태양전지의 예측 발전량과 실제 측정된 발전량을 비교하였다. 이후, 측정값을 바탕으로 두 가지 인공신경망 모델인 LSTM과 MLP에 각각 학습시켜 예측값과 측정값을 비교하였다. 이 과정에서 시계열 데이터 처리에 장점을 가진 LSTM이 예측에 더 적합한 것으로 보였다.

결과적으로, 다양한 환경 요인이 태양광 발전량 예측에 미치는 영향을 그래프로 분석하였으며, 이를 통해 예측의 정확성을 높일 수 있는 방법론을 제시하였다. 이를 통해, 향후 태양광 발전량을 더욱 정확하게 예측할 수 있을 것으로 기대된다. 이는 최적의 위치 선정 및 효율성 향상에 중요한 통찰을 제공할 것이다.

## 감사의 글

This work was supported by the International Science & Business Belt support program, through the Korea Innovation Foundation funded by the Ministry of Science and ICT. This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No.2022R111A3064285).

## 참고문헌

1. Min-Suk Kim, Yoon-Hyun Kim, Ji-Chan Kim, Yeong-Seo Park, Jong-Hwan Lee "Prediction of Solar Power Generation Using Meteorological Data", Journal of The Korean Institute of Electrical Engineers, pp. 168-169, Nov. 2023.
2. Woo Taekyeong, So Minseop, Kang Shinyoung, Shin Jongho "Development of a solar power generation prediction model utilizing photovoltaics and climate data", Korean Journal of Computational Design and Engineering, vol. 29, no. 1, pp. 33-41, Mar. 2024.
3. Da-In Hong, Jong-Hwan Lee "I-V Modeling Based on Artificial Neural Network in Anti-Reflection Coated Solar Cells", Journal of the Semiconductor & Display Technology, vol. 21, no. 3, pp. 130-134, Sep. 2022.
4. Sung-Won Choi, Ji-Hyung Ryu and Chang-Goo Lee, "Analytical Methods for the Extraction of PV panel Single-Diode model parameters from I-V Characteristic", Journal of the Korea Academia-Industrial cooperation Society, vol. 12, no. 2, pp. 847-851, Feb. 2011.
5. Sung-Won Choi, Ji-Hyung Ryu, Woo-yong Han and Chang-Goo Lee, "Extraction of PV panel model para-

- 
- meters from I-V Characteristic curve”, Institute of Control, Robotics and Systems, pp. 181-182, Dec. 2010.
6. Sera Dezso, Teodorescu Remus, Pedro Rodriguez, “PV panel model based on datasheet values”, Industrial Electronics, ISIE 2007. IEEE International Symposium on, pp. 2392-2396, Jun. 2007.
  7. Kim Kyeongmin, Kim Seongyeom, Lee Jungeun, Lee Jonghwan, “Prediction of I-V Characteristics in Si Solar Cells using Artificial Neural Network”, Journal of The Institute of Electronics and Information Engineers vol. 60, no.2, pp. 27-34, February 2023
  8. Baek Jihye, Lee Jonghwan, “Artificial Neural Network Modeling for Photovoltaic Module Under Arbitrary Environmental Conditions”, Journal of the Semiconductor & Display Technology, vol. 21, 2o. 4, pp. 135-139, Dec. 2022.
  9. G Van Houdt, C Mosquera, G Nápoles, “A review on the long short-term memory model”, Artificial Intelligence Review, pp. 3-6, May. 2020.
  10. H. Taud and J.F. Mas “Multilayer Perceptron (MLP)”, Journal of Geomatic Approaches for Modeling Land Change Scenarios, pp. 451-455, Oct. 2017.
- 
- 접수일: 2024년 8월 20일, 심사일: 2024년 9월 13일,  
게재확정일: 2024년 9월 14일