

발전량 예측 모델 기반의 태양광 모니터링 시스템 고장 예측

¹홍제성, ²박지훈, ^{*3}김영철

¹홍익대학교 소프트웨어 공학 연구실, hong@selab.hongik.ac.kr

²홍익대학교 소프트웨어 공학 연구실, jhp@selab.hongik.ac.kr

^{*3}홍익대학교 소프트웨어 공학 연구실, bob@hongik.ac.kr

Fault Prediction of Photovoltaic Monitoring System based on Power Generation Prediction Model

¹Jeseong Hong, ²Jihoon Park, ^{*3}Youngchul Kim

¹SELab. Hongik University, hong@selab.hongik.ac.kr

²SELab. Hongik University, pjh@selab.hongik.ac.kr

³SELab. Hongik University, bob@hongik.ac.kr

요약

기존의 태양광 발전 모니터링 시스템은 현재의 발전량, 과거의 발전량, 환경센서 값 등을 모니터링 한다. 이는 발전소의 효율적인 운영과 유지보수를 위한 태양광 발전량 예측이 필요하기 때문이다. 이를 위해 데이터 축적을 통해 빅데이터 기반 태양광 발전 모니터링 시스템의 발전량 예측 알고리즘 구현 방안을 제안한다. 이를 통해 미리 예측된 발전량에 비례하여, 태양광 발전 플랜트의 고장을 예측하고자 한다. 결과적으로 시스템의 고장을 예측하여 미리 점검하도록 한다.

Abstract

Existing Photovoltaic(PV) monitoring system monitors the current, past power generation, all values of environmental sensors. It is necessary to predict solar power generation for efficient operation and maintenance on the power plant. We propose a method for estimating the generation of PV data based PV monitoring system with data accumulation. Through this, we intend to find the failure prediction of the photovoltaic power plant in proportion to the predicted power generation. As a result, the administrator can predict the failure of the system it will be prepared in advance.

Keywords: Photovoltaic monitoring system, machine learning, regression, predict, neural network

1. 서론

최근 세계와 정부의 친환경 정책에 따라 신재생 에너지 설비의 보급 및 확대를 크게 늘리고 있다. 그중 태양광 에너지 발전에 대한 투자와 관심이 높아졌다. 태양광

* Corresponding Author

Received: Jun. 4, 2018, Revised: Jun. 21, 2018, Accepted: Jun. 30, 2018

에너지 발전에 필요한 단가가 줄어들고, 자원 고갈 문제가 없으며, CO₂의 방출이 낮고, 태양광 발전으로 생산된 전기를 판매하여 새로운 수익을 창출할 수 있다는 장점으로 태양광 발전은 신재생 에너지 분야에서 수요가 증가하고 많은 각광을 받고 있다. 이에 따라 태양광 발전기의 효율적인 발전과 유지보수를 위해 실시간 태양광 모니터링 시스템에 더하여 태양광 발전량 예측이 필요하다.

태양광 발전은 해가 떠있는 시간에 발전을 한다. 즉, 일사량이 많을 때 태양광 모듈에서의 발전을 한다. 이를 통하여 기상청의 지역별 기상 데이터와 태양광 발전 시스템의 데이터를 분석하여 머신 러닝을 적용해 태양광 발전량을 예측해 볼 수 있다.

본 논문에서는 머신 러닝 기법 중 1)순환 신경망(RNN)기법으로 미래의 일사량을 예측하고, 2)최소 자승 회귀 분석(OLS) 예측으로 미래의 발전량을 예측한다. 예측된 발전량과 실제 발전량을 비교하여 태양광 모듈의 유지보수를 하는 데에 기여를 한다.

2. 관련 연구

2.1. 태양광 발전 모니터링 시스템

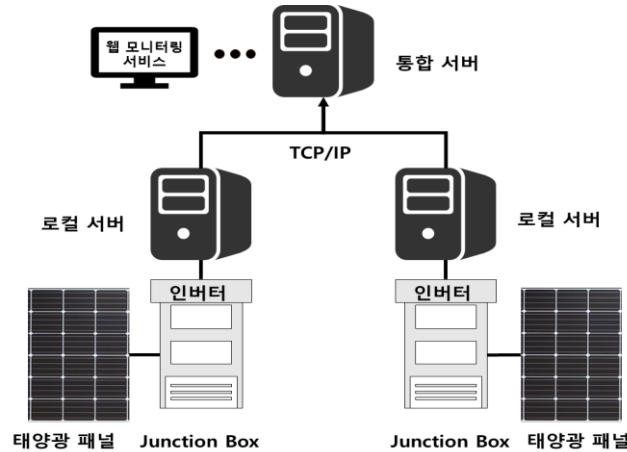


Figure 1. Structure of Integrated Photovoltaic Monitoring System

그림 1은 태양광 발전 통합 모니터링 시스템 구조이다. 태양광 발전 통합 모니터링 시스템[1]은 태양광 패널과 Junction Box, 인버터, 로컬 서버, 통합 서버가 연결되어 있다. 태양광 발전 데이터는 인버터를 통해 발전량, 출력량, 전압, 전류, 모듈 온도, 수평 일사량, 수직 일사량 등의 데이터가 각 발전소의 로컬 모니터링 서버에 저장된다. 여러 지역의 각 발전소의 로컬 서버는 축적한 데이터들을 통합 모니터링 서버로 전송한다. 통합 모니터링 서버에서는 각 로컬 서버에서 보내온 데이터를 저장하고, 데이터들을 웹 페이지에서 모니터링 할 수 있도록 데이터 정보를 웹 서비스로 제공한다.

2.2. 최소 자승 회귀 분석(OLS) 기법

OLS 회귀 분석[2]은 다중 선형 회귀 분석 기법의 종류 중 하나이다. 선형 회귀 분석[3]이란 여러 데이터가 있다면 이 데이터들을 대표할 수 있는 최적의 선형 모델(선)을 찾는 것이다. 단순 선형 회귀선을 수식으로 표현하면 $H(x) = Wx + b$ 가 된다. W 는 가중치, b 는 편차이다. $H(x)$ 의 수치에 따라 선의 모양이 변화한다. 즉, 예측에 관련되는 설명변수(독립변수) x 를 통해 예측하려는 결과 값인 반응변수(종속변수) y 를 도출해내는 것이다. 단순 선형 회귀법과 같이 다중 선형회귀 분석은 단일 설명변수가 아닌 여러 개의

설명변수를 통한 계산을 한다. 이를 수식으로 나타내면 $H(x) = W_1x_1 + W_2x_2 + W_3x_3 \dots + b$ 로 나타낼 수 있다. 그림 2는 OLS 선형회귀를 나타내는 그래프이다.

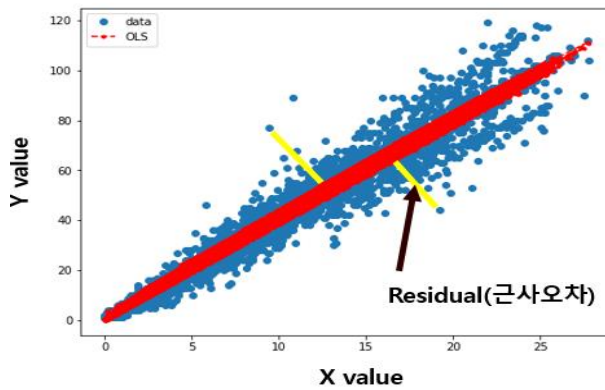


Figure 2. Ordinary Linear Regression Graph

새로운 데이터가 추정된 모델로부터 얼마나 멀리 있는 값인지를 의미하는 것을 Residual(근사오차)이라고 한다. 즉, 어떤 데이터 (x_i, y_i) 의 Residual은 $R_i = y_i - f(x_i)$ 이다. 예측 모델에 새로운 데이터를 넣은 것을 $f(x_i)$ 이고, 그 예측 결과를 y_i 로

나타낸다. 최종적으로 $\sum_{i=1}^n R_i^2 = \sum_{i=1}^n (y_i - f(x_i))^2$ 가 된다. 데이터 예측 모델의 최적의 선을 찾기 위해 Residual의 제곱의 합들을 최소화함으로써 선의 기울기를 조절한다. 이 최적의 선을 통해 새로운 데이터를 통한 결과 값을 예측한다.

2.3 순환 신경망(RNN) 기법

그림 3은 RNN[4]의 구조를 나타낸 것이다. 그림의 왼쪽은 RNN의 반복 구조를 나타낸 것이며, 오른쪽은 하나의 RNN 셀을 펼친 것이다. RNN은 입력, 망각, 출력 게이트로 구성되어 있다. 이는 인간의 뇌신경 동작을 모방하여 만든 뉴럴 네트워크 알고리즘이다[5]. 과거의 입력 값을 통해 미래에 일어날 일을 예측하는 알고리즘으로 음성인식, 언어 모델링, 번역, 이미지 처리, 주가 예측 등 시계열 데이터를 다룰 때 사용된다. RNN 셀들은 반복 루프를 돌며, 과거의 값 x_t 를 입력으로 h_t 결과 값을 출력한다. 이는 다시 h_t 값을 다음 시간의 셀에 전달을 해주어 과거의 정보를 유지한다.

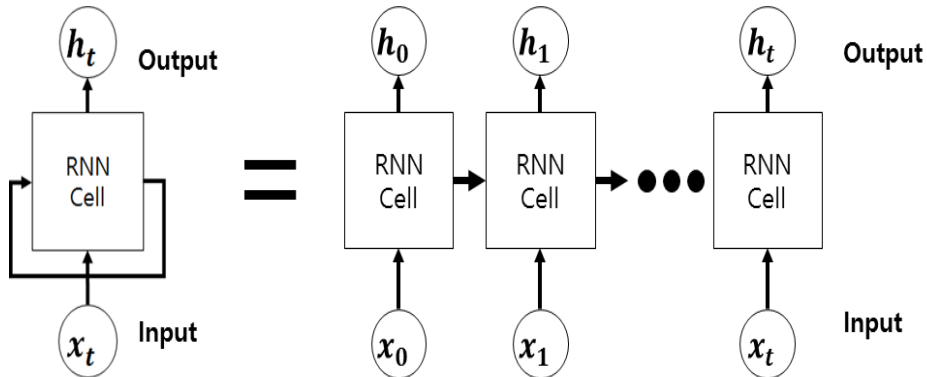


Figure 3. Basic Structure of Recurrent Neural Network Algorithm

3. 태양광 통합 모니터링 시스템 발전량 예측

3.1 데이터 상관 계수 분석

태양광 발전 시스템의 데이터는 세종시 보람초등학교 발전소의 데이터를 사용하였다. 이 데이터는 10 분 간격으로 서버에 저장된 값들이다. 표 1 은 태양광 발전 시스템의 출력 데이터와 다른 요소별 Pearson 상관 계수[6]를 측정 한 것이다. 표 1 를 통해 요소별 상관도가 높으며 의미 있는 데이터로 쓸 수 있는 것은 수평 일사량과, 수직 일사량을 알 수 있다. 일사량 데이터와 발전량 데이터 간의 상관도를 분석하기 위해 대전 지역의 1 시간 단위의 기상 데이터와 10 분 단위의 태양광 발전 시스템의 데이터를 같은 시간대로 묶는다. 그룹화를 위하여 같은 시간대의 데이터를 평균을 내는 전처리를 하였다. 결합된 데이터의 상관 계수를 측정한다. 표 2 는 기상 데이터와 태양광 시스템 데이터의 상관 계수를 측정 한 표이다. 구름이 많은 때에는 일사량이 적은 것으로 반비례 관계를 보이고, 출력과 발전량은 일사량과 비례하는 높은 상관도를 보인다.

Table 1. Correlation Coefficient Measurement table of Solar Data

DATA	현재 출력	일 발전량	출력 진류	출력 전압	입력 발전	입력 전압	수평 일사량	수직 일사량	모듈 온도
현재 출력	1.0000	0.0402	0.9978	0.5945	0.9977	0.6426	0.9843	0.9803	0.5117

Table 2. Correlation Coefficient Measurement Table between Weather Data and Solar Data

DATA	일사량	전운량	중하층 운량	현재 출력	발전량
현재 출력	0.7016	-0.4314	-0.3613	1.0000	0.9860
발전량	0.7251	-0.4281	-0.3598	0.9860	1.0000

3.2 태양광 발전량 예측 모델 알고리즘

태양광 발전량 예측에 사용될 데이터들을 선정하고, 이 데이터들을 발전량 예측 모델에 적용하여 학습시킨다.

그림 4은 태양광 통합 모니터링 시스템에 발전량 예측 모델을 적용시킨 그림이다. 태양광 모니터링 시스템 서버에 실시간으로 저장되는 태양광 통합 모니터링 시스템의 데이터와, 기상 자료 개방 포털에서 제공해주는 기상 데이터를 결합한다. 결합된 데이터에서 결측 데이터를 처리하고, 발전을 하지 않는 시간대의 데이터는 학습에 포함하지 않도록 데이터 전처리를 하고, 일정 기간 동안의 데이터를 학습시킨다.

설계한 RNN 모델은 12 시간 후의 값을 예측하도록 되어있다. 발전량 RNN 모델을 통해 발전량과 출력량 데이터를 입력하여 후의 출력량을 예측한다. 일사량 RNN 모델을 통해 중하층 운량과, 전운량, 일사량의 데이터를 입력으로 후의 일사량을 예측한다. OLS 회귀 분석 모델은 출력량과, 일사량을 설명변수로 하여 발전량을 예측하는 회귀 예측 모델이다. 두 개의 RNN 모델들로 예측한 출력량, 일사량 값을 OLS 선형 회귀 예측 모델에 넣어 다음의 발전량을 예측한다.

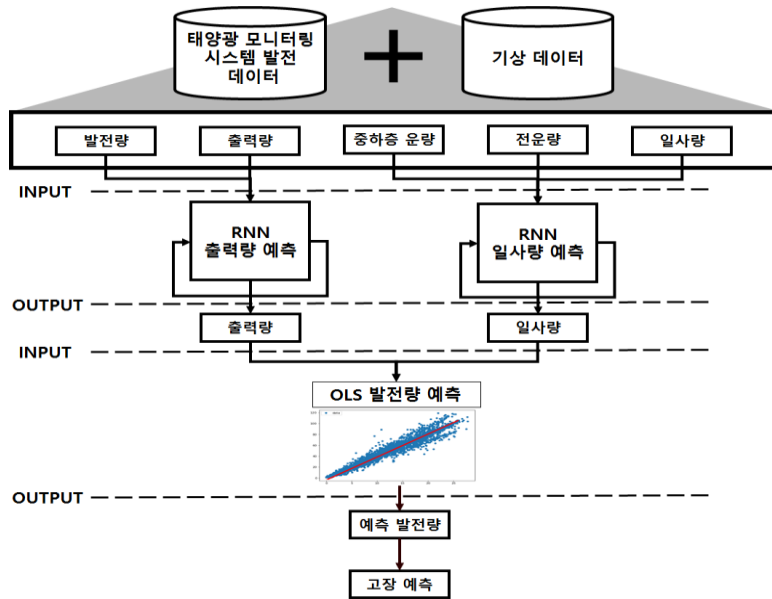


Figure 4. Photovoltaic Monitoring System with Power Generation Forecasting Model

3.3 예측 결과

그림 5는 예측 모델의 결과를 보이는 그래프이다. 그래프는 빨간 점선인 예측 값과 파란 실선으로 표현된 실제 값의 그래프를 나타낸다. 또한, 오차와 예측 평균치, 실제 평균치를 볼 수 있다. 학습시킨 데이터는 2017-01-01 ~ 2018-04-30 기간 동안의 데이터를 학습시켰다. 학습시킨 모델과 2018-05 한달의 데이터를 통하여 모델 검정을 한 결과를 그래프로 나타낸 것이다. 그 중 2018-05-21 ~ 2018-05-24 날의 결과이다. 약간의 오차가 있으나, 그래프의 곡선의 유사도는 비슷하다. 일 평균 발전량도 차이가 얼마 없이 잘 예측한 결과가 나왔다. 오차율은 매번 다른 값을 보이지만, 전체 월 단위 평균 오차율은 약 5% 수준이다.

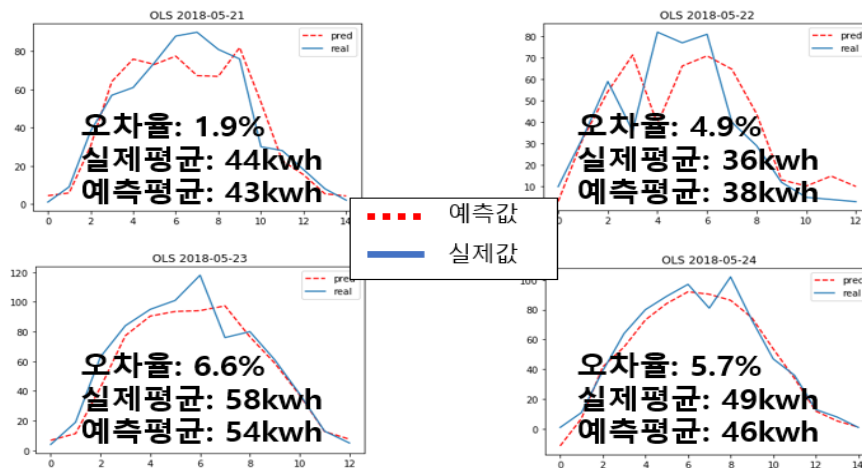


Figure 5. Power Generation Forecast Results of Sejong Boram Elementary School

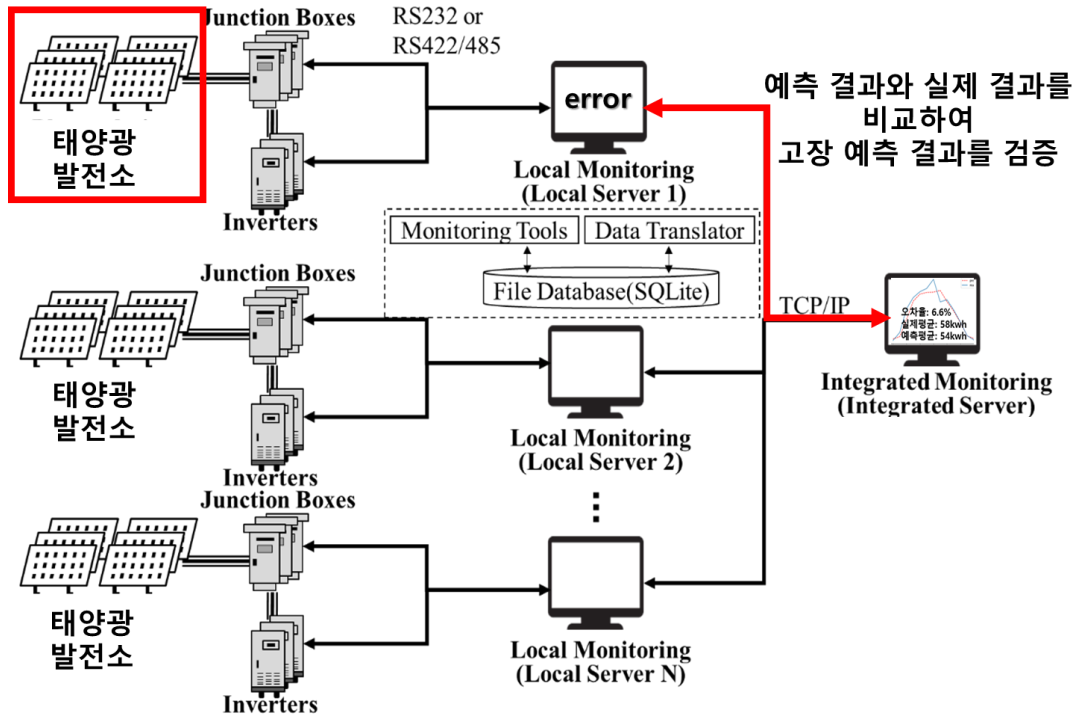


Figure 6. Error Prediction Scheme in Integrated Monitoring System

그림 6 은 전체 통합 모니터링 시스템에서 예측 결과를 검증하는 과정이다. 내일의 태양광 발전 그래프를 예측하여 그래프로 저장한다. 그리고 당일의 발전량 그래프가 오차 범위 내로 진행된다면 정상적인 발전소일 것이고 오차 범위 밖이라면 고장이 날 발전소이므로 점검이 필요하다.

4. 결론

본 논문에서는 RNN 순환 신경망 예측 모델과 OLS 선형 회귀 분석 예측 모델을 통하여 태양광 발전 시스템의 다음 날의 발전량을 예측하는 방법을 제안한다. 예측된 발전량보다 차이가 많은 수준의 발전을 지속적으로 한다면 해당 모듈의 고장이 날 확률이 높아졌다는 것을 짐작할 수 있다. 현재는 세종시에 위치한 발전소의 데이터를 사용하였으나, 세종시의 기상 데이터를 사용하지 못하고 가까운 대전시의 기상 데이터를 사용하였다. 발전소의 위치와 맞는 기상 데이터를 사용한다면 더 좋은 모델을 만들어 정확도가 높은 예측을 할 수 있을 것이다.

향후 연구에는 좀 더 정확한 일사량 예측 모델을 만들기 위해 더 많은 기상 변수를 적용할 것이다. 또한 예측 데이터와 실제 데이터 비교 분석을 통해 오차 범위를 설정하여 더 정확한 고장 예측을 할 것이다.

5. 감사의 글

이 논문은 2015 년 교육부와 한국연구재단의 지역혁신창의인력양성사업(NRF-2015H1C1A1035548)과 2017 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2017R1D1A3B03035421)과 2018 년도 정보통신산업진흥원의 정보통신, 방송 연구개발사업(개방형 OS 환경개발 및 보급, 확산 사업)의 지원을 받아 수행된 연구임(S1113-18-1001).

6. 참고문헌

- [1] Hyun Seung Son, R. Young Chul Kim, “Modeling a Photovoltaic Monitoring System based on Maintenance Perspective for New & Renewable Energy”, The 2nd International Joint Conference on Convergence, AACL 07 p.144-147, January 2016.
- [2] OLS, <https://www.statsmodels.org/dev/>
- [3] Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, “An Introduction to Statistical Learning with Applications in R”, Springer, p.59-119, 2013
- [4] RNN, <https://aikorea.org/blog/mn-tutorial-1/>
- [5] Kyunghyun Cho, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, “Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”, 2013
- [6] Aurélien Géron, (2017), hands on machine learning with scikit-learn and tensorflow, .p.85-89