

## 설명 가능한 AI를 통한 일사량 예측

조범진

22 충남 아산시 순천향로, 순천향대학교 빅데이터공학과

## Insolation Prediction with eXplainable AI

Bum-Jin Jo

22 Soonchunhyangro, Shinchang-myeon, Asan-si, Chungcheongnam-do, Korea  
Department of Big Data Engineering  
Soonchunhyang University

### Abstract

스마트 그리드의 주요 기술로 주목받는 신재생 에너지 중 태양광 발전 에너지는 영구적으로 사용 가능한 에너지이며 환경오염의 대책으로도 유용하다. 태양광 발전 시스템을 효율적으로 운용하기 위해서는 정확한 예측이 필수적이다. 인공지능 기술이 발전하여 기계학습이 상용화되고 더 나아가 심층 신경망 모델을 적용하기도 한다. 하지만 이러한 모델들을 적용해가는 과정을 이해하기 어렵고 도출되는 과정을 보여주지 않고 결과만을 보여주기도 한다. 그렇게 그 과정을 블랙박스로 표현하기도 한다. 이 블랙박스의 문제점을 해결하기 위해 설명 가능한 AI(XAI)의 기법 중 SHAP을 활용하여 설명하기 어려웠던 기계학습의 과정을 설명한다.

### I. 서론

화석연료를 사용으로 인한 환경오염 문제와 에너지의 고갈 문제가 나타나고 있다. 이를 해결하기 위해 신재생 에너지의 개발과 이용 및 보급이 이루어지고 있으며 신재생 에너지는 스마트 그리드의 주요 기술로 많은 주목을 받고 있다. 신재생 에너지 중 태양광 에너지는 영구적으로 사용 가능한 에너지이며 재생 가능한 에너지이다. 또한 에너지 조달과 사용과정에서 유해물질이나 방사선 폐기물이 나오지 않고 온실가스를 배출

하지 않아 지구 온난화에 대한 대책으로도 유용하다. 태양광 발전 시스템의 효율적인 운용을 위해서는 정확한 일사량 예측이 필요하다. 하지만 태양광 발전 시스템의 효율적인 운용을 위해서는 사전에 정확한 예측이 필수적이다. 전력 생산량에 영향을 미치는 온도, 습도, 풍속, 풍향 등과 같은 기상 정보는 기상청의 동네예보 시스템을 통해 제공되지만 미래 일사량 정보는 제공되지 않는다. 따라서 국내 여러 기관에서 비선형 관계를 효과적으로 학습 할 수 있는 기계학습(Machine Learning)을 기반으로 여러 일사량 예측 모델을 보고하였다.

최근 에너지 분야의 만족스러운 성과를 도출하는 여러 인공지능 기술을 적용하는 사례들이 증가하고 있다. 하지만 인공지능 기술의 경우 많은 부분 비선형성이 높기 때문에 결과를 도출하는 과정에 대한 이해가 직관적이지 않은 것이 문제가 된다. 이러한 특성으로 결과를 도출하는 과정을 블랙박스로 표현하기도 한다.

지도교수 : 조정기  
Tel : 041)530-1221  
E-mail : [ckcho@sch.ac.kr](mailto:ckcho@sch.ac.kr)

이에 본 논문에서는 인공지능의 도출과정 즉 블랙박스의 문제점을 해결하기 위하여 설명 가능한 AI(XAI)의 기법 중 SHAP모형을 제안하였다. SHAP의 기법은 기존의 LIME기법 등과 다르게 결과가 결정적이라는 것 외에도 이상적인 특징이 몇가지 더 있다. SHAP은 항상 결과가 일정하고 서로 영향을 미칠 가능성을 고려하며 음의 영향력 또한 고려할 수 있는 기법이다. 위 모델을 사용하여 인공지능기술을 도입할 때 설명력을 제공할 수 있다는 것을 제시한다.

## II. 관련연구

대한민국에서 가장 더운 도시라고 알려져 있는 대구 지역을 선정하여 기상청 기상자료개방포털의 종관기상

Table 1. 모델 학습을 위한 데이터 셋 구성

No.	Variable Name	Description
1	Month	Month of the year
2	Day	Day of the month
3	Hour	Hour of the day
4	Temp	Temperature
5	Humi	Humidity
6	WS	Wind speed
7	WD	Wind direction

관측자료에서 오전 8시부터 오후 6시까지 습도, 풍속, 풍향과 같이 기상청 동네예보에서 제공하는 기상 요인들과 일사량 데이터를 1시간 단위로 2011년부터 2020년까지 수집하였으며, 설명 가능한 AI(XAI)를 학습시키기 위해 Table 1 과 같이 7가지의 입력 변수를 구성하였다.

## III. 제안모형

기계학습 (Machine Learning)을 기법 중 하나인 LightGBM 활용하여 모델을 적용시켰다. 학습파라미터를 소개하면 알고리즘 타입은 dart(Dropouts meet Multiple Additive Regression Trees)로 정의하였고 learning\_rate 값은 0.01로 n\_estimators 값은 1000으로 num\_leaves 값은 64로 정의하였고 subsample은 0.5 colsample\_bytree는 1 random\_state는 42로 n\_jobs는 -1로 파라미터 값을 설정해주고 나서 모델을 학습시켰고 그것의 과정을 SHAP모델로 보여주었다. Fig. 1

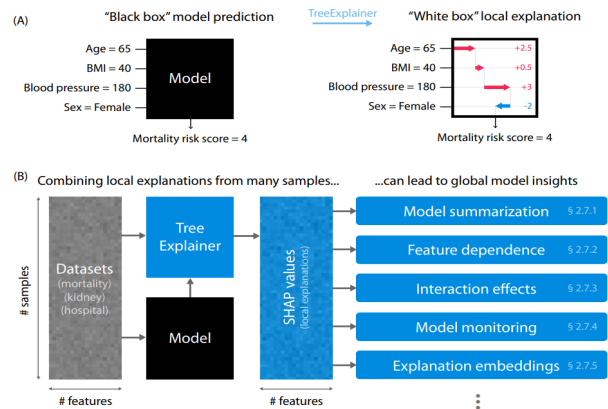


Fig. 1 설명 가능한 AI(XAI) SHAP의 과정

과 같이 블랙박스라고 불리는 모델 학습과정을 화이트 박스로 투명하게 보여주었고 Model summarization, Feature dependence, Interaction effects, Model monitoring, Explanation embeddings 등을 통해 사용자가 이해하기 쉽게 기계학습의 과정을 설명한다.

## IV. 분석결과

Fig. 2을 통해 전체 변수의 영향력이 어떠한지 쉽게 보여준다. 각 기능은 모델에 절대적인 영향을 미치기도 하는데 가장 많은 영향을 주는 것은 Hour(시간)임을 확인할 수 있으며 WD(풍향)이 가장 영향을 주지 못함을 알 수 있다.

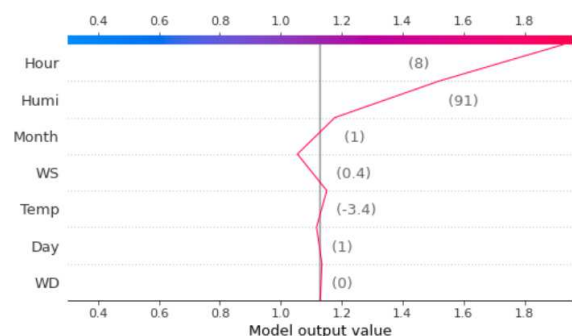


Fig. 2 model output value

또한 Fig. 3과 같이 점의 색과 분산정도를 확인하여 각 변수마다 어떤 영향을 주는지 알 수 있다. Fig. 3에서는 Humi(습도)와 Hour(시간)은 가장 큰 음의 영향을 주고 Temp(온도)는 가장 큰 양의 영향을 주고 있음을 알 수 있다.

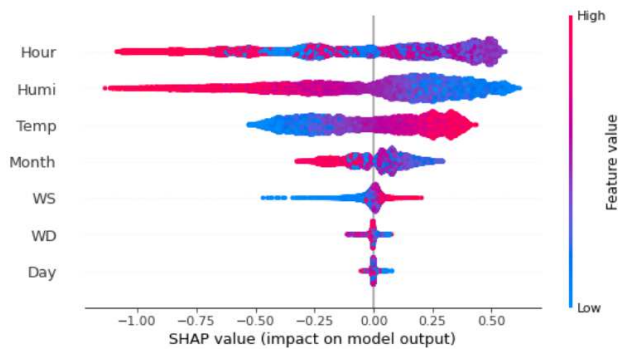


Fig. 3 summary\_plot

Fig. 4는 dependence\_plot(의존도\_플롯)은 기능 하나 하나에 대한 SHAP의 영향력 1차원 평면에 나타내어서 보여준다. 특정 데이터에 대한 SHAP의 영향력을 분해하게 됩니다. x축은 특성(=컬럼)값이고, y축은 해당 특성의 Shapley value를 가리킨다. 오른쪽 y축에 나온 특성은 현재 그린 특성과의 관계(상호작용 효과)를 나타내는데, 그래프 상에서 색깔이 수직 패턴이 나오는 경우 관계가 존재한다고 할 수 있다.

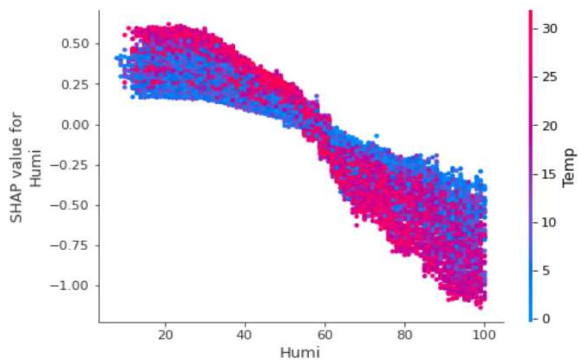


Fig. 4 dependence\_plot 1

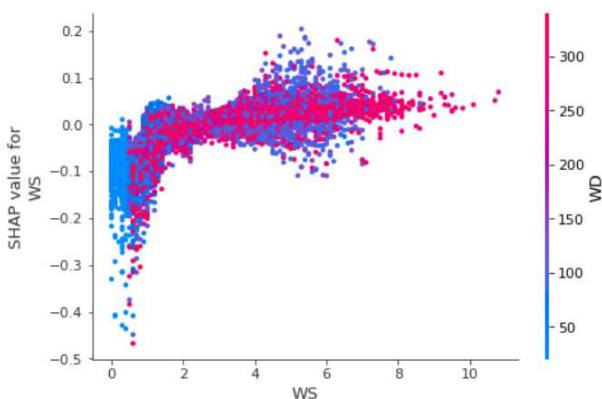


Fig. 5 dependence\_plot 2

두가지 분석자료를 대표적으로 보여주면 이와 같다. Fig. 4는 Temp(온도)와 Humi(습도)에 대한 영향력을 확인할 수 있고 Fig. 5는 WS(풍속)과 WD(풍향)에 대한 영향력을 확인할 수 있었다.

## V.결론(본인 내용 아래에 서술)

본 논문은 대한민국의 가장 더운 도시인 대구를 대상으로 일사량 예측 모델을 사용하였고 예측 모델을 선택하는과정을 설명할 수 있도록 XAL의 기법 중 SHAP 모델을 제안하였다. 예측결과와 과정을 하나하나 분석하여 어떤 이유로 예측되었는지 알 수 있었다.

향후 서울, 부산, 대구, 인천, 광주 등 다양한 지역의 일사량 데이터로 예측하여 SHAP을 적용해보고 예측 모델 또한 LightGBM 뿐만 아니라 다양한 모델들을 적용하여 각 모델들의 설명과정 또한 분석하고 가장 좋은 모델을 찾아 인공지능 분야 비전문가들에게 인공지능 모델의 이해를 도우려고 한다.

## 참고문헌

- 1) Junhak Lee, Jinwoo Jeong, Sungji Jung, Jihoon Moon and Seungmin Rho : (2022), Verification of De-Identification Techniques for Personal Information Using Tree-Based Methods with Shapley Values, Journal of Personalized Medicine, 12, 190.
- 2) Scott M. Lundberg, Gabriel Erion, Hugh Chen1, Alex DeGrave, Jordan M. Prutkin, BalaNair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee : Explainable AI for Trees: From Local Explanations to Global Understanding