

# 다양한 feature를 이용하여 태양광 발전량 예측

## 요약

최근 신재생 에너지의 중요성이 커짐에 따라 정부도 이에 대한 지원이 커지고 있다. 그 중에서 태양광 발전은 다른 신재생 에너지와 다르게 설치하기 용이하여 다양한 공기업, 사기업, 공장 등의 건물 옥상에 설치되고 있다. 현재 태양광 발전을 하는 고객과 앞으로 참여할 잠재 고객의 태양광 발전량을 예측하는 딥러닝 모델을 구현한다. 예측된 태양광 발전량은 주식회사 시너지의 태양광발전 관리 시스템에 적용할 수 있고, 이는 소규모 분산 자원 거래 시장에서 입찰 정확성을 높일 수 있고 이를 이용하여 추가 인센티브를 확보할 수 있다.

## 1. 서론

### 1.1. 연구배경

화석연료의 고갈로 인한 자원확보 경쟁 및 고유가의 지속으로 인해 신재생 에너지로의 전환은 이제 필수에 준하는 상황에 왔다. 한국을 포함해 세계적인 추세는 기존의 석탄, 석유와 더불어 원자력 발전까지 신재생 에너지로의 전환이 필요한 자원으로 보고 있으며, 각종 에너지 전환 정책들과 함께 친환경 발전을 장려하고 있다. 태양광, 태양열, 풍력 등 많은 종류의 신재생 에너지가 있지만, 현재로서는 대부분이 설치 과정에 까다로운 조건이 붙거나, 설치 비용 대비 발전효율이 떨어지고 소규모의 작업장에는 적용하기 힘든 것이 대부분이다. 태양광 발전은 다른 발전 방법에 비해 설치에 제한 사항이 적으며, 기업을 포함해 규모가 작은 일반 가정에서도 쉽게 접할 수 있다.

태양광 발전이 다른 신재생에너지에 비해 설치 조건이 까다롭지 않은 것은 맞지만, 발전의 효율을 높이기 위해서는 여러가지 요소들을 만족해야 한다. 일조량이 곧 전력생산량과 비례하기 때문에 가장 우선적으로 고려해야 하는 것이며, 패널의 온도가 너무 높거나 낮아도 발전의 효율이 떨어지며, 초기 투자비가 많이 소요된다. 이러한 요소들로 인해, 이미 운용 중인 고객 층과 잠재 고객 층 모두에게 특정 부지 (설치 후보지) 및 조건에서의 발전량 예측은 투자 대비 효율을 높일 수 있는 효과적인 방법이 될 수 있다.

현재까지 발전량 데이터와 해당 기간의 기상 정보를 딥러닝 모델에 접목한다면, 차후의 기상 정보가 있다면, 발전량을 예측할 수 있다. 또한, 태양광 발전에 영향을 미치는 요소들 각각의 가중치를 책정할

수 있고, 이를 관리 시스템에 적용해 발전 효율을 높일 수 있다. 아직 설치하지 않은 잠재고객의 경우, 설치 후보지에 일정 수준 이상의 데이터가 존재한다면 유의미한 수준의 예측이 가능할 것으로 예상되며, 초기 설치 비용의 감소와 투자 대비 효율을 높이는 파생효과를 기대할 수 있을 것이다.

## 1.2. 연구목표

충분한 기간의 시간단위 기상 측정 데이터와 발전량 데이터를 딥러닝 모델로 학습시키면, 시계열 예측이 가능하며, 이를 통해 유의미한 수준의 정확도를 가진 모델을 설계한다.

해당 모델을 통해 근미래의 날씨 예보를 받아와 이미 운용 중인 태양광 설비들의 발전량 예측 정보를 제공한다. 이 경우 직접 조작 가능한 feature의 경우 발전량을 높일 수 있는 지침을 제공할 수 있도록 한다.

또한, 설치 후보군의 features를 제공할 경우, 예상 발전량을 제공하고 해당 지역의 적합 여부를 판단한다.

## 2. 관련연구

### 2.1. 태양광 발전

현재 기존 화석 연료에서 대체에너지로 전환하는 과정에서 가장 보편적이고 소규모에도 적용할 수 있는 수단이다. 수력 발전이 80~90%, 화력 발전과 원자력 발전이 40% 가까운 효율을 보이는데 비해 태양광 발전의 발전 효율은 약 8~15%로 매우 낮다. 때문에 넓은 면적을 필요로 하며, 외부적으로 영향을 끼치는 요소들이 많다.

#### 2.1.1. 태양광 발전 원리

태양광 발전은 무한정, 무공해의 태양 에너지를 변환시켜 전기를 생산하는 발전기술이다. 태양전지를 이용한 발전방식인데, 반도체의 PN접합에 빛을 비추면 광전효과에 의해 전력이 생산되는 것을 이용한 것이다. 즉, 전자를 가지고 있는 Negative형 반도체와 양공을 가지고 있는 Positive형 반도체의 접합부에 빛이 닿으면 전자가 흐르게 되어 전류가 발생하는 원리이다.

### 2.1.2. 태양광 발전과 온도의 관계

일조시간이 길어지면 일조 표면의 온도는 점진적으로 상승하는데, 태양광 패널의 온도가 일정 수준을 초과하게 될 경우(약 25°C) 모듈마다 편차가 있지만 보통 1°C 상승마다 0.5% 정도 효율이 점점 떨어지게 되며, 패널의 수명 역시 줄어든다. 때문에 여름이 일사량이 가장 많지만, 평균 발전량은 봄, 가을, 여름, 겨울 순으로 높다. 이를 위한 해결 방안으로 온도 조절을 위해 모듈의 배치를 조정하거나 냉각수를 분사하는 등의 방법을 통해 패널의 온도를 낮추는 시도를 하고 있으며, IBM 사의 sunflower 같은 다양한 방법들이 연구되고 있다.

## 2.2. 딥러닝 모델 종류

근 몇년간, AI는 많은 관심을 받아왔고 많은 분야에서 활용되고 있다. 머신러닝, 더 깊게는 딥러닝까지 많은 연구가 이루어지고 있는데, 이 과정에서 ANN, DNN, CNN, RNN과 같은 모델들이 만들어졌다.

### 2.2.1 K-NN

K-최근접 이웃(K-Nearest Neighbor, KNN)은 지도 학습 알고리즘 중 하나이다. 어떤 데이터가 주어지면 그 주변의 데이터를 살펴본 뒤 더 많은 데이터가 포함되어 있는 범주로 분류하는 방식이다. K-NN의 특징은 훈련이 따로 필요 없다는 것이다. 훈련 데이터를 기반으로 모델을 만들고 테스트 데이터로 테스트를 하는 방식이다.

### 2.2.2 CNN

Convolutional Neural Network, 합성곱 신경망은 데이터를 feature로 추출하여 이 feature들의 패턴을 파악하는 구조이다. 크게 Convolution과 Pooling의 두 과정을 통해 진행되는데, 데이터 셋에따라 convolution layer와 pooling layer를 복합적으로 적합하게 구성하는 것이 중요하다. Convolution 과정은 데이터의 feature를 추출하는 과정이다. 데이터의 한 지점과 그 인접한 부분의 성분들을 조사하고 이를 한 장으로 만든다. Pooling 과정에서는 convolution을 거쳐서 나온 데이터의 사이즈를 줄여주는 과정을 거친다.

### 2.2.3 LSTM

RNN의 경우, 데이터간의 거리가 멀어질수록 데이터의 상관관계가 불명확해지는데 이를 개선한 모델이 바로 LSTM(Long Short Term Memory Network)이다. 가장 기본적인 RNN 모델의 경우, 학습이 거듭될수록 초기의 weight값이 유지되지 않아 학습률이 저하되는 Long-Term Dependency 현상이 생기는데 LSTM에서는 Cell State라는 layer를 추가해서 weight를 계속 기억할 것인지 결정하여 이

문제를 개선했다. 이를 통해 직전의 데이터만 참고하는 Vanilla RNN과 달리 LSTM은 cell state에서 다른 weight를 참고할 수 있게 되었다.

### **3. 프로젝트 내용**

#### **3.1. 모델 학습에 필요한 데이터**

##### **3.1.1. 태양광 발전량**

시계열 데이터로 구성된 태양광 발전량 데이터가 필요하다. 발전량 데이터는 한 시간마다 측정되어져 있다. 시계열 데이터로 구성되어 있으니 순환신경망 모델에 학습시킬 수 있다. 그리고 발전량을 측정한 날짜와 시간이 있으므로 그 시간에 맞는 다른 데이터들을 같이 사용할 수 있다.

##### **3.1.2. 날씨**

날씨는 태양광 발전에 직접적으로 영향을 미치는 데이터이다. 태양광이 얼마나 비치는지가 중요하므로 비오는 날은 거의 발전이 되지 않고, 구름의 양에 따라 발전량이 차이가 난다. 그리고 온도 역시 영향을 미치는 데이터이다. 온도가 너무 높거나 너무 낮으면 태양광 발전에 영향을 미친다는 연구가 다수 존재한다.

##### **3.1.3. 일조량**

일조량은 한 시간동안 지표면에 태양빛이 일정 밝기 이상으로 비친 시간을 측정한 데이터이다. 일사량과 착각하기 쉬운데 일사량은 한 시간동안 지표면에 전달된 태양복사에너지를 측정하는 것으로써 구름이 꺼서 흐려도 태양복사에너지는 지표면에 전달된다. 태양광 발전은 태양빛으로 발전하는 것이라 일조량이 영향을 미친다.

##### **3.1.4. 항공사진**

항공사진으로 구름의 유무를 알 수 있다. 발전소의 위치를 알 때 그 위치의 항공사진을 날짜와 시간별로 가져와서 사용할 수 있다. 항공사진도 시계열 데이터가 가능할 경우 순환신경망 모델에 사용할 수 있다.

## 3.2. 딥러닝 모델 종류

### 3.2.1. 시계열 데이터를 통한 근미래의 태양광 발전량 예측 모델

태양광 발전량, 날씨, 온도, 일조량, 항공사진 등의 feature로 사용할 데이터를 수집한다. 해당 feature 데이터들을 같은 시간에 측정한 데이터끼리 맞춰서 전처리한다. 그리고 이를 순환신경망 모델에 학습시켜서 근미래의 태양광 발전량을 예측한다.

### 3.2.2. 실제 발전소의 데이터를 바탕으로 가상 발전소의 발전량 예측 모델

수집한 날씨, 온도, 일조량, 항공사진 등의 시계열 데이터들을 발전량을 나타내도록 벡터공간에 표시하여 회귀학습을 진행하면 가상 발전소의 발전량을 예측할 수 있다. 사용할 수 있는 딥러닝 모델은 SVM 계통 모델과, k-NN 모델이 있다.

## 4. 역할 분담

장유진 - 시계열 데이터를 통한 근미래의 태양광 발전량 예측 모델

유재상 - 실제 발전소의 데이터를 바탕으로 가상 발전소의 발전량 예측 모델

## 5. 결론 및 기대효과

태양광 발전량 예측을 통해 시너지의 태양광 관리 시스템에 접목시켜 태양광 발전 효율의 증대를 도모한다. 이를 통해 전력거래 시장에서의 입찰 정확성을 올릴 수 있으며, 추가적인 이익으로 귀결된다.

태양광 설치시에 많은 지원을 해주고 있지만, 그럼에도 여전히 부담되는 가격인 것은 사실이며, 적합한 지역에 설치하지 않으면 기대만큼의 효과를 볼 수 없다. 그렇기에 부지 선정에 만전을 기해야 하는데 잠재고객에게 이 정보를 토대로 구체적이고 가시적인 발전량 예측을 제공하면 부지 선정의 정확성을 높일 수 있으며 신규 고객 유치가 용이하다.

## 6. 참고문헌

- [1] 한국에너지공단 : 신 · 재생에너지 소개 <https://www.knrec.or.kr/energy/>
- [2] 대한민국 정책브리핑 : 에너지전환 정책  
<http://www.korea.kr/special/policyCurationView.do?newsId=148864795>
- [3] 3E : Understanding high temperatures and solar power generation  
<https://www.3e.eu/heatwave-solar-production-analysis-summer-months/>