

머신 러닝을 통한 신재생에너지 예측 최적화

18009 고도형

March 19, 2020

1 서론

고객들의 필요에 따라 적절한 양의 에너지를 공급하기 위해 전력을 공급하는 회사들은 발전소의 생산량을 조절하고 다른 회사로부터 전력을 수입해서 수요를 충족시켜 왔다. 그러나 갈수록 필요한 전기에너지의 양이 증가하면서 수요가 전체 발전소의 생산량을 합한 것보다 큰 상황이 발생하는 것에 대비해 전력 시장에서 공급뿐만 아니라 수요를 조절하는 것도 필요해졌다. 그리하여 현재 전력 공급량에 맞추어 전기 사용자가 자신의 사용량을 변화시키는 수요 반응(DP; Demand Response)의 개념이 등장했다. 수요 반응은 현재 시장에서 제공 가능한 에너지의 양에 따른 인센티브를 제공하여 그에 따라 자연스럽게 소비자가 스스로의 전기 사용에 변화를 만드는 것을 말한다. 그리고 효율적인 에너지 사용을 위하여 도입된 스마트 그리드 시스템은 에너지를 절감하고 효율적으로 사용하기 위해 전력 시장에서 소비자의 능동적이고 자발적인 수요반응을 활용한다. 그리고 이를 위해 전력 수요를 상시 모니터링하며 수요의 증감에 따라 필요한 발전소의 발전량을 조절한다. 다행히 가정이나 공장의 전력 수요는 누적된 데이터를 통해 어렵지 않게 예측할 수 있고 전력을 수급하기 위한 발전소의 선택과 전력의 공급 시기를 정확히 만족시킬 수 있다.

그런데 에너지의 초과수요를 방지하고 소비자로부터 수요 반응을 이끌어내기 위해서는 전력 수요뿐만 아니라 특정 시점에서의 한계 공급량이 어느 정도인지 미리 파악할 필요가 있다. 일반적인 화력 발전이나 원자력 발전의 경우에는 시간에 따른

발전량의 기록이 크지 않아 한계 공급량을 어렵지 않게 파악할 수 있으나 대부분의 신재생에너지는 기상 상황에 따라 발전량의 차이가 커 예측이 불가능하고 관리하기가 쉽지 않다. 이러한 신재생에너지의 발전량은 기상 파라미터 값과 깊은 연관성을 가지는 만큼 정확한 예측을 위해 주로 기상 데이터를 이용한 기계적 모델이 채택되고 있다.

2 선행 연구

Emil Isaksson과 Mikael Karpe Conde는 논문 Solar Power Forecasting with Machine Learning Techniques에서 머신 러닝 기반의 태양광 발전량 예측 모델을 설계하였다. 이 모델은 지도학습 알고리즘 중 하나인 K-최근접 이웃 방법을 활용하여 높은 정확도로 한계 공급량을 예측하였다. 연구진은 과거의 각종 기상 데이터들을 바탕으로 해당 환경 아래에서 실제 발전량과 근접하게 예측할 수 있는 모델을 훈련시키고 실제로 기상 예보 데이터를 접목시켜 미래의 전력 생산량을 예측하고자 하였다.

3 K-Nearest Neighbor

연구에 사용된 KNN 알고리즘은 K-최근접 이웃(K-nearest neighbor)이라 불리는 지도학습 알고리즘이다. 본래 데이터를 여러 개의 클래스로 구별하는 분류(Classification) 문제에 적합하게 설계되었으나 논문에서는 이를 회귀 문제에 적용하였다. 기본적으로 주어진 데이터 X 로부터 특정 클래스에 속한 주변 데이터의 위치를 고려하여 가장 높은 확률을 가지는 클래스로 Y 의 값을 추정한다. 즉, 다음과 같은 과정을 거친다. 우선, 어떠한 테스트 데이터 x_0 와 클래스 j 가 존재할 때, x_0 로부터 가장 가까운 유클리드 거리(Euclidean distance)를 가지는 K 개의 데이터들의 순열 N_0 을 구한다.

$$dist(x_0, K) = \sqrt{\sum_{i=1}^K (x_0 - X_i)^2} \quad (1)$$

다음으로, 추정값 $Y = j$ 를 만족할 확률을 x_0 로부터 가장 가까운 K 개의 데이터 중 클래스 j 에 속하는 것의 개수와 전체 K 의 값을 바탕으로 아래와 같이 정의한다.

$$Pr(Y = j|X = x_0) = \frac{1}{K} \sum_{i_0} I(y_i = j) \quad (2)$$

회귀 문제에 적용되는 알고리즘은 데이터의 추정값을 이웃한 데이터 값들의 평균으로 정의하여 동작한다.