

태양광 발전량 예측

Random Forest Regressor 의 하이퍼파라미터 변경에 따른 성능결과 분석

18011547 컴퓨터공학과 최예린

제공된 데이터 셋

가상데이터를 이용하여 태양광 발전량을 예측하기 위한 데이터 셋이다.

| | |
|--------|-----------|
| Hours | 시간 |
| Minute | 분 |
| DHI | 수평면 산란일사량 |
| DNI | 직달일사량 |
| WS | 풍속 |
| RH | 상대습도 |
| T | 기온 |
| Target | 태양광 발전량 |

Random Forest Regressor - hyperparameter 값에 따른 성능분석

이 레포트에서는 n_estimator 와 min_sample_split, max_depth 의 변경에 따른 성능결과를 분석합니다.

| no | parameter | value | score |
|----|------------------|---------|---------|
| 1 | - | default | 0.99853 |
| 2 | n_estimator | 15 | 0.99819 |
| 3 | n_estimator | 20 | 0.99818 |
| 4 | n_estimator | 30 | 0.99819 |
| 5 | n_estimator | 50 | 0.99824 |
| 6 | min_sample_split | 3 | 0.99852 |
| 7 | min_sample_split | 4 | 0.99851 |
| 8 | min_sample_split | 5 | 0.99847 |
| 9 | max_depth | 10 | 0.99746 |
| 10 | max_depth | 15 | 0.99849 |
| 11 | max_depth | 20 | 0.99852 |
| 12 | max_depth | 25 | 0.99853 |
| 13 | max_depth | 30 | 0.99853 |

이 데이터 셋에서는

n_estimator default 값인 10 일 때 가장 좋은 성능을 보였다.

min_sample_split default 값인 2 일 때 가장 좋은 성능을 보였으며, 값이 커질수록 성능은 줄어들었다.

max_depth default 값이 None 인 파라미터이다. 트리의 최대 깊이를 결정하는 것인데 25 이후로는 default 로 설정했을 때와 같은 성능을 보이는 것으로 확인된다.