

회귀 모델 성능 평가 지표(MAE, MSE, RMSE, MAPE 등)

회귀 모델(regression model)을 평가할 때 주로 MAE, MSE, RMSE, MAPE 등을 사용한다.

MAE(Mean Absolute Error) = 평균 절대 오차

실제 정답 값과 예측 값의 차이를 절댓값으로 변환한 뒤 합산하여 평균을 구한다.

특이값이 많은 경우에 주로 사용된다.

값이 낮을수록 좋다.

$$MAE = \frac{\sum |y - \hat{y}|}{n}$$

장점

- 직관적임
- 정답 및 예측 값과 같은 단위를 가짐

단점

- 실제 정답보다 낮게 예측했는지, 높게 했는지를 파악하기 힘들
- 스케일 의존적임(scal dependency): 모델마다 예류 크기가 동일해도 에러율은 동일하지 않음

MSE(Mean Squared Error) = 평균 제곱 오차

실제 정답 값과 예측 값의 차이를 제곱한 뒤 평균을 구한다.

값이 낮을수록 좋다.

$$MSE = \frac{\sum_{i=1}^n (y - \hat{y})^2}{n}$$

장점

- 직관적임

단점

- 제곱하기 때문에 1미만의 에러는 작아지고, 그 이상의 에러는 커짐
- 실제 정답보다 낮게 예측했는지, 높게 했는지를 파악하기 힘들
- 스케일 의존적임(scal dependency): 모델마다 예류 크기가 동일해도 에러율은 동일하지 않음

RMSE(Root Mean Squared Error) = 평균 제곱근 오차

MSE에 루트는 씌워서 에러를 제곱해서 생기는 값의 왜곡이 줄어든다.
값이 낮을수록 좋다.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n}}$$

장점

- 직관적임

단점

- 제곱하기 때문에 1미만의 에러는 작아지고, 그 이상의 에러는 커짐
- 실제 정답보다 낮게 예측했는지, 높게 했는지를 파악하기 힘들
- 스케일 의존적임(scal dependency): 모델마다 예류 크기가 동일해도 에러율은 동일하지 않음

MAPE(Mean Absolute Percentage Error) = 평균 절대 비율 오차

MAE를 비율, 퍼센트로 표현하여 스케인 의존적 에러의 문제점을 개선한다.
값이 낮을수록 좋다.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y - \hat{y}}{y} \right|$$

장점

- 직관적임
- 다른 모델과 에러율 비교가 쉬움

단점

- 실제 정답보다 낮게 예측했는지, 높게 했는지를 파악하기 힘들
- 실제 정답이 1보다작을 경우,무한대의 값으로 수렴할 수 있음

기타

그 외에 MPE, R2 score 등이 있다.

MPE(Mean Percentage Error)

MAPE에서 절대값을 제외하여 계산한다.

모델이 underperformance 인지 overperformance인지 판단할 수 있다.(음수이면 overperformance, 양수이면 underperformance)

$$MPE = \frac{100}{n} \sum_{i=1}^n \frac{y - \hat{y}}{y}$$

R2 score = R squared

다른 지표(MAE, MSE, RMSE)들은 모델마다 값이 다르기 때문에 절대 값만 보고 성능을 판단하기 어렵다.

R2 score는 상대적인 성능을 나타내기 비교가 쉽다.

실제 값의 분산 대비 예측값의 분산 비율을 의미한다.

1에 가까울 수록 좋다.

$$R^2 = 1 - \frac{SSE}{SST}$$

$$SSE = \frac{1}{n} \sum_{i=1}^n (y^i - \hat{y}^i)^2 = MSE$$

$$SSE = \frac{1}{n} \sum_{i=1}^n (y^i - \bar{y})^2$$

- $\hat{y}$: y 의 예측값
- $\bar{y}$: y 의 평균 값