

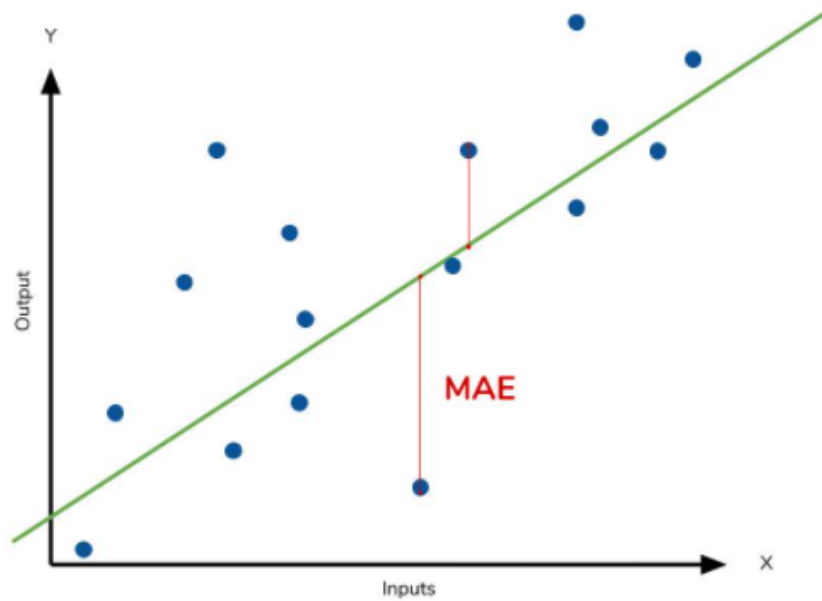
● MAE(Mean Absolute Error)

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

모델의 "예측값과 실제값의 차이"를 모두 더한다는 개념

✧ 절대값을 취하기 때문에 가장 직관적으로 알 수 있는 지표이다.

- MSE 보다 특이치에 robust한다. (robust란 ? 이상치에 대한 저항도 가지고 있고, 데이터 특성을 잘 나타냄)
- 절대값을 취하기 때문에 모델이 underperformance 인지 overperformance 인지 알 수 없다.
  - underperformance: 모델이 실제보다 낮은 값으로 예측
  - overperformance: 모델이 실제보다 높은 값으로 예측

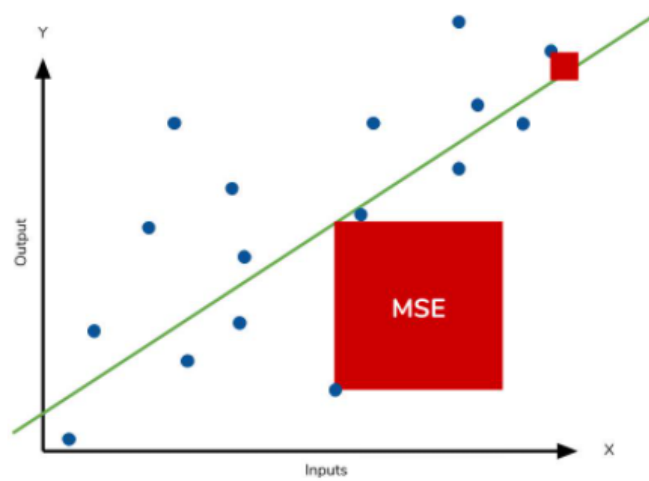


● **MSE**(Mean Squared Error)

$$MSE = \frac{\sum (y - \hat{y})^2}{n}$$

★ 제곱을 하기 때문에 (MAE와는 다르게) 모델의 예측값과 실제값 차이의 면적의 합이다. 이런 차이로 특이값이 존재하면 수치가 많이 늘어난다.

- 특이치에 민감하다



## Root Mean Squared Error(RMSE)

RMSE는 MSE에 루트를 씌워 다음과 같이 정의합니다.

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

RMSE를 사용하면 오류 지표를 실제 값과 유사한 단위로 다시 변환하여 해석을 쉽게 합니다.

outlier를 무시함.

①

**MAE:** <sup>1</sup> you can use this when you have very few or no outliers in your data or in a better way <sup>2</sup> when you want to ignore the outliers while fitting your model to your data.

**MSE/RMSE:** you use this when you have a large number of outliers in your data and want to accommodate them while fitting your model.

outlier를 고려함, 대부분의 데이터셋에 outlier가 포함되어 있기 때문에,  
MAE보다 MSE를 더 많이 사용함.

Hi, you use mean squared error when you want to penalize more bigger errors versus smaller errors, this because the square of a big number is way bigger than the one of a smaller number :)