Regression Metrics

1. **Mean Absolute Error (MAE) :-**

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Advantages - • same unit as target
• robust to outliers

Disadvantages - • not differentiable

2. **Mean Squared Error (MSE) :-**

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

Advantages - • differentiable (can be used as cost function)

Disadvantages - • squared unit of target
• not robust to outliers

3. **Root Mean Squared Error (RMSE) :-**

$$RMSE = \sqrt{MSE}$$

Advantages - • same unit as target

Disadvantage - • not as robust as MAE (to outliers)

4. R-squared (R2 score) :-

R2 score (also called Coefficient of Determination) is used to evaluate the performance of regression models. It is the amount of variation in the target variable which is explained by the input variables.

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

where, SSR → sum of squares of residuals
SST → Total sum of squares

- R2 score of baseline model is 0.
- Best possible R2 score is 1 (when predicted value = actual value)
- R2 score can be negative.

5. Adjusted R-squared :-

When an irrelevant feature is added to a dataset, the R2 score increases instead of decreasing. To avoid this, adjusted R-squared is used.

$$R_a^2 = 1 - \left[\left(\frac{n-1}{n-k-1}\right) \times (1-R^2)\right]$$

where, n → no. of input columns

k → " " independent input columns

R^2 → R squared

R_a^2 → adjusted R squared