

Para saber mais

No último vídeo aprendemos a avaliar um modelo de machine learning com o “RMSE” ou “root mean squared error”. Essa é uma forma muito conhecida para avaliar modelos de regressão, no entanto existem outras formas de avaliar esse tipo de modelo.

Mean Absolute Error

O erro médio absoluto, é a média da diferença entre os valores observados e previstos. De certa forma ele nos indica o quão longe as previsões foram do valor real, mas não indica a direção do erro, que também é conhecido como under prediction e over prediction.

A representação matemática se dá por:

$$MeanAbsoluteError = \frac{1}{N} \sum_{j=1}^N |y_j - \hat{y}_j|$$

R^2

O coeficiente de determinação, ou R quadrado, é outra técnica que pode ser empregada para avaliar modelos de regressão.

Uma vantagem importante nessa métrica é o fato dela ser livre de escala. Se ao avaliarmos um modelo, o RMSE for 70, não saberemos se esse valor é bom ou ruim sem entender a necessidade de negócio e de avaliar outras métricas, como o desvio padrão, média, etc. No caso do R^2 , um valor próximo de 1, indica um modelo com aproximadamente 0 erro.

A formula matemática do RMSE é:

$$R^2 = 1 - \frac{MSE(model)}{MSE(baseline)}$$

Onde o MSE(model) é igual ao Mean Squared Error do modelo que aprendemos a calcular nessa aula, ou simplesmente o RMSE².

Já o MSE(baseline) é calculado com a seguinte fórmula:

$$MSE(baseline) = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2$$