

Reflexão sobre Avaliação de Modelos de Regressão

O Erro Absoluto Médio (*Mean Absolute Error*)

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

[01] Em que situações o MAE é útil?

1. Quando precisamos de uma métrica que seja representativa para os dados que estamos avaliando (ex.: quando prevendo o preço de uma casa, um MAE de 10.000 pode significar que em média nós erramos 10 mil reais)
2. Quando queremos uma métrica para medir o quanto estamos errando (no intervalo $[0, \infty[$)

[02] Como se dá a interpretação dos resultados no MAE?

O MAE vai de zero até o infinito. Quanto menor esse valor melhor

O MAE dá a nós a diferença média entre as observações (valores verdadeiro) e os resultados do nosso modelo (valores preditos). Um ponto positivo do MAE é que o seu resultado sempre está na mesma escala de medida que da amostra, então se estamos tentando prevendo o preço de uma casa, um MAE de 10.000 pode significar que em média nós erramos 10 mil reais - repare que a unidade de medida é a mesma que o preço de uma casa.

[03] O MAE pode ser considerado análogo a alguma medida de tendência central?

O MAE assemelha-se a média, pois nos diz o quão longe em média estamos da medida correta

[04] Como o MAE se comporta com *outliers*?

O MAE dá o mesmo peso para os outliers do que para todos os outros exemplos da amostra.

O Percentual do Erro Médio Absoluto (*Mean Absolute Percentage Error*)

$$\text{MAPE}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}|}{\max(\epsilon, |y_i|)}$$

[05] Em que situações o MAPE seria útil?

Quando queremos analisar o quanto erramos em percentual em relação ao valor que estamos tentando prever.

Exemplo: Se estamos tentando prever o valor de uma casa que custa 10.000 (true value) e prevemos 12.000 (predicted value), o nosso MAPE é 0.2 (ou 20%).\

Essa análise é especialmente útil no mundo dos negócios, onde é de muito valor compararmos sempre um resultado em relação a um baseline (patamar base)

[06] Como é a interpretação dos resultados no MAPE?

Em geral, valores menores que 20% são desejáveis e menores que 10% são muito bons.

Fonte: <https://stephenallwright.com/good-mape-score/>

[07] Como o MAPE se comporta com *outliers*?

Assim como o MAE, o MAPE não é nem resistente nem muito sensível a outliers. Ele dá para os outliers o mesmo peso que qualquer outro exemplo da amostra.

O Erro Quadrático Médio (*Mean Squared Error*)

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

[08] Em que situações o MSE é útil?

1. Quando queremos dar um peso maior aos outliers
2. Quando precisamos de uma métrica para função custo de um modelo (pois o MSE gera apenas um só mínimo global e local, aka bowl shaped function)

[09] Como é a interpretação dos resultados no MSE?

Se o MSE é 0, significa que nosso estimador prevê os dados com precisão perfeita (o ideal, mas normalmente não é possível). Quanto maior o valor do MSE, mais distante são as previsões do estimador dos valores reais

[10] O MSE pode ser análogo a alguma medida de variabilidade?

O MAE assemelha-se a variância, pois nos diz o quão longe em média estamos da medida correta e assim como a variância a diferença entre o valor da nossa instância e o valor que estamos tentando medir é elevado ao quadrado.

[11] Como o MSE se comporta com *outliers*?

Quando há muitos outliers no dataset isso pode afetar drasticamente o valor do MSE, pois o MSE é muito sensível a outliers

A Raiz Quadrada do Erro Quadrático Médio (*Root Mean Squared Error*)

$$RMSE(y, \hat{y}) = \sqrt{MSE(y, \hat{y})}$$

[12] Em que situações o RMSE é útil?

Como no RMSE os erros são elevados ao quadrado antes da média, o RMSE dá um peso relativamente alto para erros grandes. Isso significa que o RMSE é útil quando grandes erros são particularmente indesejáveis.

Um caminho para sabermos se o RMSE é uma boa métrica é uma pergunta como:

O RMSE tem o benefício de penalizar mais erros grandes, então pode ser mais apropriado em alguns casos, por exemplo, se *estar errado por 10 é mais que duas vezes pior do que estar errado por 5, então provavelmente o RMSE é uma boa métrica de performance*. Mas se estar errado por 10 é apenas duas vezes pior que sendo desligado por 5, então MAE é mais apropriado.

Referência:

[13] Como é a interpretação dos resultados no RMSE?

Quanto mais próximo o RMSE estiver de 0, mais preciso será o modelo. Mas como RMSE é retornado na mesma escala que o alvo para o qual você está prevendo não há uma regra geral sobre como interpretar seu intervalos de valores. A interpretação do seu valor só pode ser avaliada dentro do seu conjunto de dados

[14] O RMSE pode ser análogo a alguma medida de variabilidade?

O RMSE assemelha-se ao desvio padrão, pois assim como o desvio padrão é a raiz quadrada da variância, o RMSE é a raiz quadrada do MSE

[15] Como o RMSE se comporta com *outliers*?

O RMSE penaliza erros grandes mais do que erros pequenos, pois ele é derivado de uma métrica que eleva ao quadrado os erros.

O Coeficiente de Determinação: R^2

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

[16] Em que situações o R^2 é útil?

Quando queremos saber o quão bem o nosso modelo está performando. O R^2 é especialmente útil porque ele está numa escala fixa (0-1)

[17] Como é a interpretação dos resultados no R^2 ?

O R^2 apresenta valores que variam entre 0 e 1; quanto mais próximo de 1 mais um nosso modelo consegue representar a variância dos nossos dados.

[18] Como o R^2 se comporta com *outliers*?

Ele é muito sensível a outliers

O Erro Absoluto Mediano (*Median Absolute Error*)

$$MedAE(y, \hat{y}) = mediana(|y_1 - \hat{y}_1|, \dots, |y_n - \hat{y}_n|)$$

[19] Em que situações o MedAE é útil?

Quando precisamos de uma métrica mais insensível para outliers

[20] Como se dá a interpretação dos resultados no MedAE?

Para uma amostra suficiente grande e com não muita variância, o MedAE representa uma instância dessa amostra que representa o quanto, em valores absolutos, uma amostra típica está errando.

[21] O MedAE pode ser considerado análogo a alguma medida de tendência central?

MedAE é análogo a mediana, pois utiliza a mediana para representar uma instância em si da amostra para compor uma experiência típica

[22] Como o MedAE se comporta com *outliers*?

Uma vez que a amostra seja suficientemente grande e sua variância não seja muito alta, o MedAE é muito pouco sensível a outliers, pois se baseia na mediana do módulo dos erros.