

Métricas de Desempenho - Regressão

Como medir o desempenho de modelos de regressão?

Diferentemente de ferramentas para avaliar problemas de classificação, casos em que deve-se prever a categoria de um data point desconhecido, os modelos de regressão preveem uma variável numérica contínua do tipo razão ou intervalar.

O Erro Absoluto Médio (*Mean Absolute Error*)

O MAE é uma medida de erro que computa a média aritmética dos erros absolutos.

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

```
In [1]: from metricas import mae
```

```
In [2]: y = [13, -0.5, 2, 7]
        y_pred = [2.5, 0.0, 2, 8]

        mae(y, y_pred)
```

```
Out[2]: 3.0
```

QUESTÕES PARA REFLEXÃO

[01] Em que situações o MAE é útil?

Resposta: Pelo fato de não elevar as diferenças ao quadrado, essa medida torna-se uma opção não tão ideal para lidar com problemas delicados. Contudo, é uma métrica sólida para modelos que devem prever muitos dados ou dados sazonais, como em previsões de números de casos de doenças, nas quais prever a tendência e sazonalidade dos números é mais importante do que os valores absolutos de cada dia.

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

Resposta: Essa medida apresenta valor mínimo 0 e não apresenta valor máximo. Quanto mais próximo de 0 mais precisas são as previsões, medindo apenas a distância do valor real, independente de ser acima ou abaixo.

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

Resposta: Sim, já que consiste na **média** das distâncias entre valores preditos e reais.

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

Resposta: Diferentemente do MSE e do RMSE, essa métrica não "pune" tão severamente os outliers do modelo.

O Percentual do Erro Médio Absoluto (*Mean Absolute*

Percentage Error)

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

```
In [1]: from metricas import mape

y = [0, 2, 3, 4]
y_pred = [1, 2, 3, 8]

mape(y, y_pred)
```

Out[1]: 2.5e+19

QUESTÕES PARA REFLEXÃO

[05] Em que situações o MAPE seria útil?

Resposta: O MAPE é uma métrica útil por ser extremamente intuitiva, tanto para a interpretação do programador, quanto para a comunicação de resultados com pessoas sem conhecimento técnico e não lida tão bem se tratando de problemas com um grande alcance de números, como uma regressão que prevê uma variável que vai de 10 a 20.000. Geralmente é usada em relatórios de gerenciamento, porque o erro é medido como uma porcentagem e assim, é possível fazer comparações entre erros percentuais do modelo entre produtos.

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

Resposta: A medida se dá em porcentagem e assim como o MSE e o MAE, quanto menor o valor, mais preciso seria o modelo de regressão. Por exemplo, ter um MAPE=10% significa que, em média, nosso modelo faz previsões que erram por 10% do valor real.

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

Resposta: Assim como o MAE, essa métrica não “pune” tão severamente os outliers do modelo.

O Erro Quadrático Médio (*Mean Squared Error*)

O MSE é uma medida de erro que computa a média dos quadrados dos erros.

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

```
In [4]: from metricas import mse
```

```
In [5]: y = [13, -0.5, 2, 7]
y_pred = [2.5, 0.0, 2, 8]

mse(y, y_pred)
```

Out[5]: 27.875

QUESTÕES PARA REFLEXÃO

[08] Em que situações o MSE é útil?

Resposta: Dado que essa métrica eleva o erro ao quadrado, previsões muito distantes do real aumentam o valor da medida muito facilmente, o que a torna uma métrica de avaliação excelente para problemas nos quais grandes erros não são tolerados, como é o caso de exames médicos e projeções de preços.

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

Resposta: Quanto maior, pior é o modelo.

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

Resposta: Sim, análogo à variância.

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

Resposta: Essa métrica, assim como o RMSE, "pune" severamente os grandes erros.

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

O RMSE é computado pela raiz quadrada do MSE. Representa uma medida de variabilidade dos acertos de um modelo.

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

```
In [2]: from metricas import rmse
```

```
In [3]: y = [13, -0.5, 2, 7]
y_pred = [2.5, 0.0, 2, 8]

rmse(y, y_pred)
```

```
Out[3]: 5.2796780204857185
```

QUESTÕES PARA REFLEXÃO

[12] Em que situações o RMSE é útil?

Resposta: Tendo em vista essa diferença de unidades, o RMSE entra como uma forma de melhorar a interpretabilidade da métrica, acertando a unidade. Lida muito bem com problemas em que grandes erros não são tolerados e com uma melhor interpretabilidade. Frequentemente usada em séries temporais.

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

Resposta: Assim como o MSE, quanto maior for seu número pior é o modelo.

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

Resposta: Sim, como se trata da raiz quadrada do MSE pode ser análogo ao desvio padrão.

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

Resposta: Por conta do expoente ao quadrado que o erro assume, essa métrica é bastante sensível a outliers (valores discrepantes) e, caso tenha muitos erros significativos em sua análise, essa métrica poderá ser extrapolada.

O Coeficiente de Determinação: R^2

O coeficiente de determinação representa a proporção da variância (de y) que é explicada pelas variáveis independentes do modelo. O R^2 fornece uma indicação da qualidade do ajuste e portanto uma medida do quão bem instâncias não vistas serão previstas pelo modelo.

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

```
In [6]: from metricas import r2
```

```
In [7]: # Predição perfeita
# =====

y = [1, 2, 3, 4]
y_pred = [1, 2, 3, 4]

r2(y, y_pred)
```

```
Out[7]: 1.0
```

```
In [8]: # Predição imperfeita
# =====

y = [3, -0.5, 2, 7]
y_pred = [2.5, 0.0, 2, 8]

r2(y, y_pred)
```

```
Out[8]: 0.9486081370449679
```

O R^2 é muito sensível a *outliers*.

```
In [9]: # Outliers
# =====

y = [13, -0.5, 2, 7] # 3 => 13
y_pred = [2.5, 0.0, 2, 8]

r2(y, y_pred)
```

```
Out[9]: -0.04510837727006445
```

QUESTÕES PARA REFLEXÃO

[16] Em que situações o R^2 é útil?

Resposta: Muito utilizada em modelos de regressões da área de finanças, o R-Quadrado, ou Coeficiente de Determinação, é uma métrica que visa expressar a quantidade da variância dos dados que é explicada pelo modelo construído. Essa métrica, apesar de conseguir identificar algumas relações lineares entre o modelo de regressão e os dados.

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

Resposta: O valor do seu R-Quadrado varia de 0 a 1 e geralmente é representado em porcentagem. Por exemplo, um $R^2 = 75\%$ nos diz que 75% da variância de nossos dados podem ser explicados pelo modelo construído, enquanto os outros 25%, teoricamente, se tratariam de uma variância residual.

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

Resposta: É muito sensível aos outliers.

O Erro Absoluto Mediano (*Median Absolute Error*)

O MedAE é calculado ao obter a mediana das diferenças absolutas entre o alvo e a predição.

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

```
In [10]: from metricas import medae

y = [13, -0.5, 2, 7]
y_pred = [2.5, 0.0, 2, 8]

medae(y, y_pred)
```

```
Out[10]: 0.75
```

[19] Em que situações o MedAE é útil?

Resposta: É útil basicamente porque é essencialmente insensível a outliers (desde que não haja muitos deles). Isso ocorre porque é a mediana de todos os valores absolutos dos resíduos, e a mediana não é afetada pelos valores nas caudas. Com certeza será muito mais útil se combinados com outras métricas.

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

Resposta: Como a métrica é uma mediana, então podemos interpretar da seguinte forma: o valor encontrado nos diz que metade das predições estão entre 0 e o valor obtido com o MedAE e a outra metade está acima desse valor.

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

Resposta: Sim, como diz em seu nome o MedAE é análogo à mediana.

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

Resposta: É menos sensível ou robusto a alguns outliers.