

Formação Cientista de Dados

Estatística II: Métricas de Erros



Métricas de Erros

- Previsão de valores numéricos (reais, inteiros)
- Métricas diferentes da previsão de categorias

- Uso:
 - Regressão linear
 - Regressão ML
 - Series Temporais
 - Etc.



Quando uma previsão é feita...



- Existe uma diferença entre a previsão e o que ocorreu...
- Temos que medir esta diferença!
 - Para saber a qualidade do nosso modelo!
 - Para podermos melhorá-lo
 - Para podermos fazer benchmarks

Como eu sei?



Diferença...

Previsto	Realizado
3,34	3,00
4,18	4,00
3,00	3,00
2,99	3,00
4,51	4,50
5,18	4,00
8,18	4,50



Mean Erro (ME)

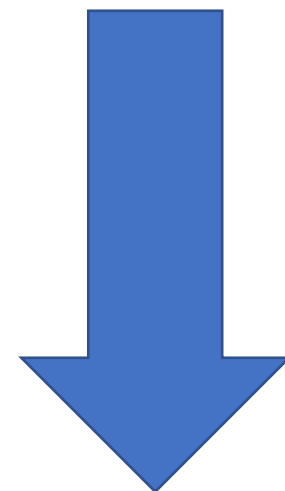
Dependente de Escala

- A média da diferença entre realizado e previsto

Previsto	Realizado	Dif.
3,34	3,00	-0,34
4,18	4,00	-0,18
3,00	3,00	0
2,99	3,00	0,01
4,51	4,50	-0,01
5,18	4,00	-1,18
8,18	4,50	-3,68
		-5,38

$$ME = \frac{\sum_{i=1}^n y_i - x_i}{n}$$

$$ME = \frac{-5,38}{7} = -0,76$$



Mean Absolute Error (MAE)

Dependente de Escala

- A média da diferença absoluta entre o realizado e o previsto

Previsto	Realizado	Dif. Absoluta
3,34	3,00	0,34
4,18	4,00	0,18
3,00	3,00	0
2,99	3,00	0,01
4,51	4,50	0,01
5,18	4,00	1,18
8,18	4,50	3,68
		5,4

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

$$\text{MAE} = \frac{5,4}{7} = 0,77$$



Root Mean Squared Error (RMSE)

Depende de Escala

- O desvio padrão da amostra da diferença entre o previsto e o teste

Previsto	Realizado	Dif. ao Quad.
3,34	3,00	0,1156
4,18	4,00	0,0324
3,00	3,00	0
2,99	3,00	1E-04
4,51	4,50	1E-04
5,18	4,00	1,3924
8,18	4,50	13,5424
		15,083

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^T (\hat{y}_t - y_t)^2}{T}}$$

$$\text{RMSE} = \sqrt{\frac{15,083}{7}}$$

$$\text{RMSE} = 1,46$$



Mean Percentage Error (MPE)

Independente de Escala (%)

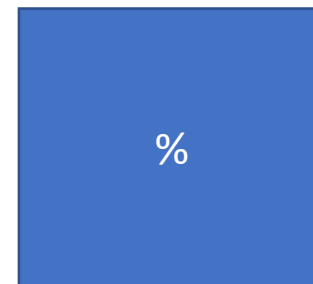
- Diferença percentual de erro

Previsto	Realizado	Erro %
3,34	3,00	-11,3333
4,18	4,00	-4,5
3,00	3,00	0
2,99	3,00	0,333333
4,51	4,50	-0,22222
5,18	4,00	-29,5
8,18	4,50	-81,7778
		-127

$$MPE = \frac{100\%}{n} \sum_{t=1}^n \frac{a_t - f_t}{a_t}$$

$$MPE = \frac{-127}{7}$$

$$MPE = -18,14$$



Mean Absolute Percentage Error (MAPE)

Independente de Escala (%)

- Diferença absoluta percentual de erro

Previsto	Realizado	Erro abs.	Erro % abs.
3,34	3,00	0,1156	0,1133333
4,18	4,00	0,0324	0,045
3,00	3,00	0	0
2,99	3,00	1E-04	0,0033333
4,51	4,50	1E-04	0,0022222
5,18	4,00	1,3924	0,295
8,18	4,50	13,5424	0,8177778
			1,2766667

$$\text{MAPE} = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

$$\text{MAPE} = \frac{1,2766667}{7}$$

$$\text{MAPE} = 0,18$$

%

Com Funciona...

	ME	RMSE	MAE	MPE	MAPE
Test set	-0.1333333	0.9565215	0.8133333	-0.5505051	9.994949

	ME	RMSE	MAE	MPE	MAPE
Test set	0.4	0.6389575	0.6	9.036341	10.70301



Calcule: Root Mean Squared Error (RMSE)

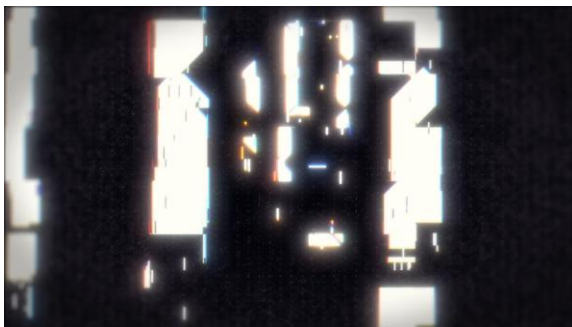


Previsto	Realizado	
3,12	3,00	0,0144
6,18	7,20	1,0404
12,30	11,00	1,69
		2,7448

$$RMSE = \sqrt{\frac{\sum_{t=1}^T (\hat{y}_t - y_t)^2}{T}}$$

$$RMSE = \sqrt{\frac{2,7448}{3}}$$

$$RMSE = 0,9565215$$



Planilha

Previsto	Realizado	Diferença	Dif. Abs.	Dif. Quad.	Erro %	Erro % abs
3,34	3	-0,34	0,34	0,1156	-11,3333	11,33333
4,18	4	-0,18	0,18	0,0324	-4,5	4,5
3	3	0	0	0	0	0
2,99	3	0,01	0,01	1E-04	0,33333	0,333333
4,51	4,5	-0,01	0,01	1E-04	-0,22222	0,222222
5,18	4	-1,18	1,18	1,3924	-29,5	29,5
8,18	4,5	-3,68	3,68	13,5424	-81,7778	81,7778

ME	-0,76857
MAE	0,77143
RMSE	1,46789
MPE	-18,1429
MAPE	18,2381