

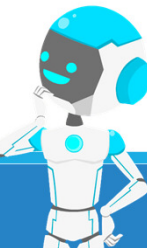
Cenário

Buscando modelo com melhor desempenho em projeto de Machine Learning. No primeiro teste:

- Modelo Naive Bayes tem desempenho de 78%
- Modelo Multilayer Perceptron tem desempenho de 80%

Porém, o custo Computacional é o seguinte:

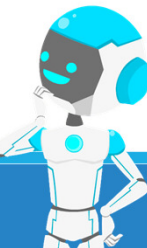
- Naive Bayes: criação do modelo em 3 minutos
- Multilayer Perceptron: criação do modelo em 2 horas



Pergunta:

Existe diferença estatisticamente significativa entre os modelos (Naive Bayes: 78% e Multilayer Perceptron: 80%), ou esta diferença é devido ao acaso?

Se estes 2% forem reais, usaremos Multilayer Perceptron. Caso contrário, usaremos Naive Bayes



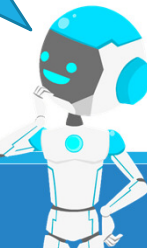
Solução: Teste de Hipótese Comparando Duas Médias Populacionais

Primeiro passo, coletar amostras aleatórias dos dois casos, ou seja, criar vários modelos de Naive Bayes e de Multilayer Perceptron, e anotar os resultados. Vamos supor que faremos 40 testes com cada algoritmo (a tabela abaixo é o resultado parcial):

Naive Bayes	Multilayer Perceptron
78	80
78	79
78	82
79	79
80	80
79	79
82	78
82	80
78	79
78	82
...	...

Observação:

Você precisa usar um processo aleatório de geração de dados de teste e treino para gerar os modelos, caso contrário todos terão o mesmo resultado!



Teste de Hipótese comparando Duas Médias Populacionais

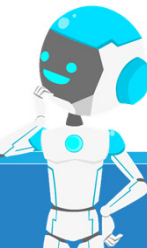
Considerando x (Naive Bayes) e y (Multilayer Perceptron)

Etapas:

Calcular as médias (\bar{X} e \bar{Y}) e os desvios padrão (s_x , s_y) das amostras

Anotar o tamanho das amostras (40)

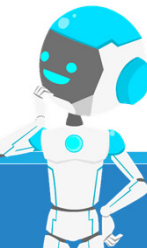
- $\bar{X} = 79,2$
- $\bar{Y} = 79,8$
- $s_x = 2,61$
- $s_y = 2,31$
- $n_1 = 40$
- $n_2 = 40$



Teste de Hipótese comparando Duas Médias Populacionais

A hipótese nula é que a diferença da média das duas populações é igual a zero:

$$H_0 = \mu_x - \mu_y = 0$$



Teste de Hipótese comparando Duas Médias Populacionais

Estatística de Teste

$$\frac{(\bar{X} - \bar{Y})}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}}$$

- $\bar{X} = 79,2$
- $\bar{Y} = 79,8$
- $s_x = 2,61$
- $s_y = 2,31$
- $n_1 = 40$
- $n_2 = 40$

Solução:

Calcula a diferença entre as médias as amostras: $(79,2 - 79,8) = -0,6$

Calcula o erro padrão: $\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}} = \sqrt{\frac{6,81}{40} + \frac{5,33}{40}} = 0,54$

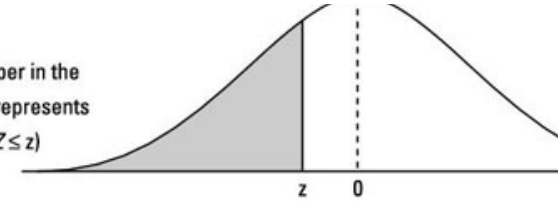
Divisão medias pelo erro padrão: $\frac{-0,6}{0,54} = -1.11$



Teste de Hipótese comparando Médias Populacionais

Procurar – 1.11 na tabela $Z = 0,1335$

Number in the
table represents
 $P(Z \leq z)$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949

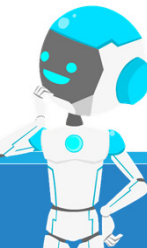
Teste de Hipótese comparando Duas Médias Populacionais

$$p = 0,1335$$

$$\text{Teste não igual (bicaudal): } 2 \times p = 0,1335 = 0,267$$

$$\text{Alfa} = 0,05$$

$$\text{Valor-p } 0,267 > 0,05$$



Teste de Hipótese Comparando Duas Médias Populacionais

Conclusão

- Não há evidências para rejeitar H_0
- Não existe uma diferença estatística significativa entre Naive Bayes e Multilayer Perceptron

