

Lista 3

Questão 1:

	TVP	TFN	TFP	TVN	Precisão	Recall	F-measure
A	30/60	30/60	25/160	135/160	30/55	30/60	0,5217
B	60/70	10/70	20/150	130/150	60/80	60/70	0,8000
C	60/90	30/90	25/130	105/130	60/85	60/90	1,2918

TVP: verdadeiro positivo. Tudo que era X e foi classificado como X dividido por todo mundo que era X no total.

TFN: falso negativo. Tudo que era X e foi classificado como não sendo X dividido por todo mundo que era X.

TFP: falso positivo. Tudo que não era X e foi classificado como X dividido por todo mundo que não é X.

TVN: verdadeiro negativo. Tudo que não era X e foi classificado como não sendo X dividido por todo mundo que não era X.

Precisão: tudo que acertou no atributo. Acerto de X dividido pela soma total da coluna de X.

Recall: tudo que acertou na classe. Acerto de X dividido pela soma total da linha de X.

F-measure: $2 \times \text{precisão} \times \text{recall} / (\text{precisão} + \text{recall})$

Questão 2:

Tanto a precisão quanto o recall de B foram os maiores, sendo assim, pode-se afirmar que a classe B obteve mais acertos do que as classes A e C.

Questão 3:

Jogar	Aparência			Temperatura			Umidade		Ventando	
	Sol	Nublado	Chuva	Quente	Agradável	Fria	Alta	Normal	Sim	Não
SIM 9/14	2/9	4/9	3/9	2/9	4/9	3/9	3/9	6/9	3/9	6/9
NÃO 5/14	3/5	0/5	2/5	2/5	2/5	1/5	4/5	1/5	3/5	2/5

$\text{SIM} = 9/14 \times 3/9 \times 3/9 \times 6/9 \times 3/9 = 0,01587301587 = 82,24\%$

$\text{NÃO} = 5/14 \times 2/5 \times 1/5 \times 1/5 \times 3/5 = 0,003428571429 = 17,76\%$

Total = 0,0193015873

Questão 4:

Lendo e tratando arquivos:

<https://colab.research.google.com/drive/18FvmUrTyleV95-FIINpGpgrYMUOiLw-e?usp=sharing>

Naive Bayes:

<https://colab.research.google.com/drive/171oZyAfqE54Sqj95ilINDQbXJGxGyl-0?usp=sharing>

Questão 5:

Random Forest:

<https://colab.research.google.com/drive/1YbAJ-bdDaXnYcpdNZJfkYPM2rs15m2fF?usp=sharing>