

4ª Sessão prática – problemas

1 - Explique em poucas palavras como funciona a estimação de erro por camadas e diga justificando se ela tem uma tendência otimista ou pessimista.

2 - Sabendo que se pretende efetuar imputação por “atribuição do valor médio dentro da classe” qual a classe do exemplo V1=2,6, V2=3,7 e V3=2,8 se for utilizado um classificador *k-nearest neighbour* com k=3 que use a distância *City-block* ?

V1	V2	V3	Classe
3,8	2,8	6,3	1
?	9,0	5,6	2
6,1	7,3	8,4	1
?	4,6	0,3	1
1,3	9,0	?	2
0,5	1,0	3,3	2

3 – Considere a tabela de probabilidades seguinte (assuma uma distribuição Gaussiana para o S3):

	P(Doença)	P(S1 D)	P(S2 D)	P(S3 D)
Doença A	22	83	11	M =10 V=2
Doença B	68	90	37	M=12 V=3
Doença C	10	56	70	M=8 V=5

Tendo em conta a tabela de probabilidades apresentada qual a probabilidade da doença B para um doente que declare não apresentar o sintoma S1 e apresentar S3=9,5?

$$P(x=v \mid c) = \frac{1}{\sqrt{2\pi}\sigma_c^2} e^{-\frac{(v-u_c)^2}{2\sigma_c^2}}$$

4 – Tenha em conta a tabela de probabilidades que se apresenta seguidamente. Sabendo que o paciente se fez acompanhar dos resultados de uma análise que aumentaram a probabilidade da doença D1 em 10% (sem dar qualquer informação sobre as restantes), qual a probabilidade de o paciente não ter a doença D1 se este declarar não ter o sintoma S3? Utilize a precisão necessária para que o resultado apresente um erro aceitável.

	P(Doença)	P(S1 D)	P(S2 D)	P(S3 D)
D1	20	15	95	80
D2	45	75	15	20
D3	35	80	30	25

$$P(D \mid S) = \frac{P(S \mid D) \times P(D)}{P(S \mid D) \times P(D) + P(S \mid \text{not } D) \times P(\text{not } D)}$$