Exame de 1ª Época de Percepção
2003 / 2004

1. Considere o problema de classificação em duas classes, usando uma característica, com as seguintes densidades condicionadas: $P(x \mid w_1) \sim N(0,1)$, $P(x \mid w_2) \sim N(1,4)$

Assuma $P(w_1) = P(w_2) = \frac{1}{2}$ e custos 0-1 nas decisões.

(a) **[0.5 valores]** Esboce as duas distribuições num mesmo gráfico.

(b) **[2.5 valores]** Derive o limiar de decisão de Bayes, mostrando no gráfico de (a) as regiões de decisão associadas com as classes w_1 e w_2.

(c) **[1 valor]** Suponha que as probabilidades a priori se alteram para: $P(w_1) = 0.6$ e $P(w_2) = 0.4$. De que forma se altera o limiar de decisão calculado na alínea (b)? Desenhe as novas regiões de decisão.

(d) **[1 valor]** Mantenha agora $P(w_1) = P(w_2) = \frac{1}{2}$, mas altere os custos de decisão para

$c = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$, em que cij é o custo de decidir classe w_i quando a classe verdadeira é w_j.

Calcule as novas regiões de decisão.

2. [3 valores] Dado um conjunto de observações independentes, $X = \{x_1, \ldots, x_n\}$, com distribuição $P(x_i) = \frac{1}{2b} e^{-\frac{|x_i - a|}{b}}$, em que *a* é conhecido e *b* é um parâmetro desconhecido, determine uma expressão para o estimador de máxima verosimilhança de *b*.

3. Considere o conjunto de 5 amostras representadas na tabela seguinte usando duas características, x e y:

	x	y
1	4	4
2	8	4
3	15	8
4	24	4
5	24	12

(i) [3 valores] Construa um dendograma admitindo que a distância entre clusters é definida por $D(C_i, C_j) = \max_{a \in C_i, b \in C_j} d(a,b)$, e que *d*(.,.) corresponde à distância Euclideana. Expliciteas matrizes das distâncias entre clusters ao longo do processo de construção do dendrograma.

(ii) [1 valor] Utilize o dendograma para particionar os dados. Justifique o critério de escolha do número de *clusters*.

4. [3 valores] Considere uma rede neuronal com entrada x e saída escalar y, cujo comportamento depende de um vector de pesos w. Sabendo que $y = f(\sum_i w_i x_i)$, em que f é uma função conhecida, deduza um algoritmo de treino em tempo real por forma a minimizar o risco empírico $R_e = \sum_{x \in X} (d(x) - y(x))^2$.

5. [1 valor] Desenhe uma rede neuronal que classifica uma amostra como pertencente à classe 1 se essa amostra produz um valor positivo para $D = 34 + 8x_1 - 7x_2 + x_3$, classificando a amostra na classe 0 caso o valor resultante para D seja negativo. Indique, justificando a sua resposta, a topologia da rede, a função de activação das unidades e os valores dos pesos.

6. A sequência de observações $O = 2132$ foi gerada pelo seguinte modelo de Markov não observável:

$$A = \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix},\ B = \begin{bmatrix} 0.2 & 0.4 & 0.2 & 0.2 \\ 0.3 & 0.1 & 0.1 & 0.5 \end{bmatrix},\ \pi = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}$$

Determine:

(a) **[2 valores]** A sequência de variáveis de estado mais provável para a sequência de observações indicada acima.

(b) **[2 valores]** Sabendo que no instante t=2 o processo se encontra no estado 2, determine a probabilidade da sequência de observações *O*.