

Examen del Bloc 2 de Sistemes Intel·ligents (tipus B)

ETSINF, UPV, 18 de desembre de 2017. Puntuació: nencerts - nerrors/3.

1 [B] Quina de les següents expressions és *incorrecta*?

- A) $\sum_x P(x | y) = 1, \forall y$
- B) $\sum_y P(x | y) = 1, \forall x$
- C) $\sum_x \sum_y P(x, y) = 1$
- D) $\sum_x P(x | u) = \sum_y P(y | w), \forall u, w$

2 [B] Es tenen dos magatzems de taronges: 1 i 2. El 65% de les taronges es troben al magatzem 1 i la resta al 2. Se sap que al magatzem 1 hi ha un 1% de taronges no aptes per al consum; i un 3% al 2. Suposeu que es distribueix una taronja no apta per al consum. Quina és la probabilitat P que siga del magatzem 1?

- A) $0.00 \leq P < 0.25$
- B) $0.25 \leq P < 0.50$ $P = P(A = 1 | C = 0) = \frac{P(A=1)P(C=0|A=1)}{P(C=0)} = \frac{P(A=1)P(C=0|A=1)}{P(A=1)P(C=0|A=1) + P(A=2)P(C=0|A=2)} = 0.38$
- C) $0.50 \leq P < 0.75$
- D) $0.75 \leq P$

3 [D] Siga $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ un objecte donat mitjançant una seqüència de N vectors de característiques, el qual es vol classificar en una C de classes. Indica quin dels següents classificadors *sí* és d'error mínim (\mathbf{x}_2^N denota $\mathbf{x}_2, \dots, \mathbf{x}_N$):

- A) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(\mathbf{x}_1 | c) p(\mathbf{x}_2^N | \mathbf{x}_1)$
- B) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(\mathbf{x}_1 | c) p(\mathbf{x}_2^N | \mathbf{x}_1, c)$
- C) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(\mathbf{x}_1, c) p(\mathbf{x}_2^N | \mathbf{x}_1)$
- D) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(\mathbf{x}_1, c) p(\mathbf{x}_2^N | \mathbf{x}_1, c)$

4 [C] Siga un classificador en 3 classes per a $\mathbf{x} = (x_1, x_2)^t \in [0, 1]^2$ amb les distribucions de probabilitat donades a la dreta. Quina és la probabilitat d'error p_e del classificador?

- A) $0.65 \leq p_e$
- B) $0.45 \leq p_e < 0.65$
- C) $0.35 \leq p_e < 0.45$ $.1 \cdot 0 + .2 \cdot .02 + .3 \cdot .5 + .4 \cdot 2/3 = .42$
- D) $p_e < 0.35$

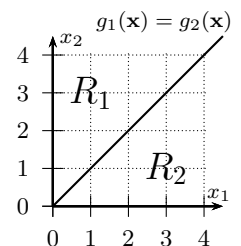
x_1	x_2	$p(c = 1 \mathbf{x})$	$p(c = 2 \mathbf{x})$	$p(c = 3 \mathbf{x})$	$p(\mathbf{x})$
0	0	1.0	0.0	0.0	0.1
0	1	0.01	0.01	0.98	0.2
1	0	0.25	0.5	0.25	0.3
1	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0.4

5 [C] Siga un problema de classificació en quatre classes d'objectes a \mathbb{R}^3 . Es té un classificador de funcions discriminants lineals amb vectors de pesos (en notació homogènia): $\mathbf{w}_1 = (-2, 1, 2, 0)^t$, $\mathbf{w}_2 = (0, 2, 2, 0)^t$, $\mathbf{w}_3 = (1, 1, 1, 0)^t$ i $\mathbf{w}_4 = (3, 0, 0, 1)^t$. Indica a quina classe s'assignarà l'objecte $\mathbf{x} = (1, 2, 2)^t$ (no en notació homogènia).

- A) 4. $3 + 0 \cdot 1 + 0 \cdot 2 + 1 \cdot 2 = 5$
- B) 3. $1 + 1 \cdot 1 + 1 \cdot 2 + 0 \cdot 2 = 4$
- C) 2. $0 + 2 \cdot 1 + 2 \cdot 2 + 0 \cdot 2 = 6$
- D) 1. $-2 + 1 \cdot 1 + 2 \cdot 2 + 0 \cdot 2 = 3$

6 [D] En la figura es representen frontera i regions de decisió d'un classificador binari. Quin dels següents parells de vectors de pesos correspon al classificador de la figura?

- A) $\mathbf{w}_1 = (-1, -1, -2)^t$ i $\mathbf{w}_2 = (-1, -2, -1)^t$ $x_2 = x_1$ $R_1 : x_2 < x_1$ $R_2 : x_2 > x_1$
- B) $\mathbf{w}_1 = (1, -1, -2)^t$ i $\mathbf{w}_2 = (0, -2, -1)^t$ $x_2 = x_1 + 1$ $R_1 : x_2 < x_1 + 1$ $R_2 : x_2 > x_1 + 1$
- C) $\mathbf{w}_1 = (-1, 1, 2)^t$ i $\mathbf{w}_2 = (0, 2, 1)^t$ $x_2 = x_1 + 1$ $R_1 : x_2 > x_1 + 1$ $R_2 : x_2 < x_1 + 1$
- D) $\mathbf{w}_1 = (1, 1, 2)^t$ i $\mathbf{w}_2 = (1, 2, 1)^t$ $x_2 = x_1$ $R_1 : x_2 > x_1$ $R_2 : x_2 < x_1$



7 [D] Siga un problema de classificació en 3 classes, $c = 1, 2, 3$, per a objectes representats mitjançant vectors de característiques bidimensionals. Es tenen 3 mostres d'entrenament representades en notació homogènia: $\mathbf{x}_1 = (1, 1, 2)^t$ de la classe $c_1 = 1$, $\mathbf{x}_2 = (1, 2, 3)^t$ de la classe $c_2 = 2$ i $\mathbf{x}_3 = (1, 3, 1)^t$ de la classe $c_3 = 3$. Així mateix, es té un classificador lineal definit pels vectors de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (2, -8, 0)^t$, $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-5, -2, -1)^t$ i $\mathbf{w}_3 = (w_{30}, w_{31}, w_{32}) = (-2, 1, -10)^t$. Si apliquem una iteració de l'algorisme Perceptró a partir d'aquests vectors de pesos, amb factor d'aprenentatge $\alpha = 1$ i marge $b = 1.5$, llavors:

- A) Es modificaran els vectors de pesos \mathbf{w}_1 i \mathbf{w}_2 .
- B) Es modificaran els vectors de pesos \mathbf{w}_1 i \mathbf{w}_3 .
- C) Es modificaran els vectors de pesos \mathbf{w}_2 i \mathbf{w}_3 .
- D) No es modificarà cap vector de pesos.

- 8 **B** En el procés d'entrenament d'un arbre de classificació, un node intern t té un grau d'impuresa $\mathcal{I}(t) > 0$. Un dels "splits" produeix un decrement d'impuresa igual a $\mathcal{I}(t)$. Indica l'afirmació correcta:
- No és possible aconseguir aqueix decrement d'impuresa.
 - Aquest "split" genera dos nodes purs.
 - Aquest "split" genera un node pur i un altre impur.
 - Aquest "split" genera dos nodes impurs.
- 9 **B** Per a un problema de classificació de dades bidimensionals $\mathbf{x} = (x_1, x_2)$ en dues classes disposem d'un arbre de classificació. Quin tipus de fronteres de decisió defineix el node arrel?
- $a \cdot x_1 + b \cdot x_2 + c = 0$ on $a \neq 0 \wedge b \neq 0$
 - $a \cdot x_1 + b \cdot x_2 + c = 0$ on $a = 0 \vee b = 0$ **Fronteres paral·leles als eixos**
 - $a \cdot x_1 + b \cdot x_2 + c = 0$ on $a \neq 0 \vee b = 0$
 - $a \cdot x_1 + b \cdot x_2 + c = 0$ on $a = 0 \vee b \neq 0$
- 10 **D** Supposeu que estem aplicant l'algorisme d'aprenentatge d'arbres de classificació per a un problema de quatre classes: 1, 2, 3 i 4. L'algorisme ha aconseguit un node t que inclou una dada de cada classe, açò és, 4 en total. Es pretén avaluar la qualitat d'una partició del node t mitjançant un "split" $s = (j, r)$, que divideix les dades en dos nodes t_1 i t_2 de la següent forma: les dades de les classes 1 i 2 queden en el node t_1 i les dades de les classes 3 i 4 queden en el node t_2 . El decrement d'impuresa $\Delta\mathcal{I}(j, r, t)$ (mesurat com entropia) per a quantificar la qualitat d'aquesta partició és:
- $\Delta\mathcal{I}(j, r, t) < 0.0$.
 - $0.0 \leq \Delta\mathcal{I}(j, r, t) < 0.5$.
 - $0.5 \leq \Delta\mathcal{I}(j, r, t) < 1.0$.
 - $1.0 \leq \Delta\mathcal{I}(j, r, t)$. $\Delta\mathcal{I}(j, r, t) = \mathcal{I}(t) - \hat{P}_t(t_1)\mathcal{I}(t_1) - \hat{P}_t(t_2)\mathcal{I}(t_2) = 2 - \frac{1}{2} \cdot 1 - \frac{1}{2} \cdot 1 = 1$
- 11 **B** Indica quina de les següents afirmacions sobre un arbre de classificació construït mitjançant l'algorisme d'aprenentatge d'arbres és *incorrecta*.
- En cada node t la suma per a totes les classes de $P(c | t)$ és 1.
 - En cada node t , la probabilitat a posteriori de qualsevol classe c , $P(c | t)$, és sempre major o igual que el menor dels pesos o probabilitats de decisió dels seus dos fills.
 - La impuresa d'un node, mesurada com entropia, no pot ser menor que 0 ni major que $\log_2 C$, on C és el nombre de classes.
 - Si N és el nombre de dades d'aprenentatge, la profunditat de l'arbre no serà major que N encara que, en la pràctica, sol ser proporcional a $\log_2 N$.
- 12 **A** En la figura de la dreta es representen 4 mostres bidimensionals. Quin és el nombre de clústers que minimitza la suma d'errors quadràtics per a les dites 4 mostres?
- 4 $J = 0$
 - 3
 - 2
 - 1
- 13 **A** La figura a la dreta mostra una partició de 4 punts bidimensionals en 2 clústers (representats mitjançant els símbols \bullet i \circ). La transferència del punt $(2, 3)^t$ del clúster \bullet al \circ condueix a una variació de la SEQ, ΔJ , tal que:
- $-1 \geq \Delta J$. $\Delta J = \frac{1}{2} - \frac{3}{2} = -1$
 - $-\frac{1}{2} \geq \Delta J > -1$.
 - $0 \geq \Delta J > -\frac{1}{2}$.
 - $\Delta J > 0$.
- 14 **B** En la figura de la dreta es mostra una partició de 4 punts bidimensionals de 2 clústers. La transferència del punt $(1, 1)^t$ del clúster \bullet al clúster \circ
- produeix un decrement en la SEQ.
 - produeix un increment en la SEQ.
 - no altera la SEQ.
 - produeix una SEQ negativa.
- 15 **B** Considereu l'algorisme C -mitjanes de Duda i Hart. Indiqueu quina de les següents afirmacions és *correcta*:
- Quan un clúster es queda buit, aquest clúster s'elimina.
 - La seua bona eficàcia computacional s'aconsegueix gràcies al càlcul incremental de la variació de distorsió i dels vectors mitjana de clúster.
 - Determina el nombre de clústers que minimitza la suma d'errors quadràtics (SEQ).
 - Cap de les anteriors.

