Examen del Bloque 2 de Sistemas Inteligentes (tipo A)

ETSINF, UPV, 18 de diciembre de 2017. Puntuación: num_aciertos - num_errores/3.

1 $\boxed{\text{A}}$ ¿Cuál de las siguientes expresiones es *incorrecta*?

A) $\sum_y P(x \mid y) \;=\; 1,\ \forall x$

B) $\sum_x P(x \mid y) \;=\; 1,\ \forall y$

C) $\sum_x \sum_y P(x, y) \;=\; 1$

D) $\sum_x P(x \mid u) \;=\; \sum_y P(y \mid w),\ \forall u, w$

2 $\boxed{\text{B}}$ Se tienen dos almacenes de naranjas: 1 y 2. El 65 % de las naranjas se hallan en el almacén 1 y el resto en el 2. Se sabe que en el almacén 1 hay un 1 % de naranjas no aptas para el consumo; y un 3 % en el 2. Supóngase que se distribuye una naranja no apta para el consumo. ¿Cuál es la probabilidad P de que provenga del almacén 1?

A) $0.00 \leq P < 0.25$

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

C) $0.50 \leq P < 0.75$

D) $0.75 \leq P$

3 $\boxed{\text{D}}$ Sea $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ un objeto dado mediante una secuencia de N vectores de características, el cual se quiere clasificar en una de C clases. Indica cuál de los siguientes clasificadores *sí* es de error mínimo ($\mathbf{x}_2^N$ denota $\mathbf{x}_2, \ldots, \mathbf{x}_N$):

A) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1 \mid c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1)$

B) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1, c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1)$

C) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1 \mid c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1, c)$

D) $c(\mathbf{x}) = \underset{c=1,\ldots,C}{\arg\max}\ p(\mathbf{x}_1, c)\, p(\mathbf{x}_2^N \mid \mathbf{x}_1, c)$

4 $\boxed{\text{B}}$ Sea un clasificador en 3 clases para $\mathbf{x} = (x_1, x_2)^t \in [0,1]^2$ con las distribuciones de probabilidad dadas a la derecha. ¿Cuál es la probabilidad de error p_e del clasificador?

A) $p_e < 0.35$

B) $0.35 \leq p_e < 0.45$ $\quad .1 \cdot 0 + .2 \cdot .02 + .3 \cdot .5 + .4 \cdot 2/3 = .42$

C) $0.45 \leq p_e < 0.65$

D) $0.65 \leq p_e$

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 $\boxed{\text{B}}$ Sea un problema de clasificación en cuatro clases de objetos en $\mathbb{R}^3$. Se tiene un clasificador de funciones discriminantes lineales con vectores de pesos (en notación homogénea): $\mathbf{w}_1 = (-2, 1, 2, 0)^t$, $\mathbf{w}_2 = (0, 2, 2, 0)^t$, $\mathbf{w}_3 = (1, 1, 1, 0)^t$ y $\mathbf{w}_4 = (3, 0, 0, 1)^t$. Indica a qué clase se asignará el objeto $\mathbf{x} = (1, 2, 2)^t$ (*no* en notación homógenea).

A) 1. $\quad -2 + 1 \cdot 1 + 2 \cdot 2 + 0 \cdot 2 = 3$

B) 2. $\quad 0 + 2 \cdot 1 + 2 \cdot 2 + 0 \cdot 2 = 6$

C) 3. $\quad 1 + 1 \cdot 1 + 1 \cdot 2 + 0 \cdot 2 = 4$

D) 4. $\quad 3 + 0 \cdot 1 + 0 \cdot 2 + 1 \cdot 2 = 5$

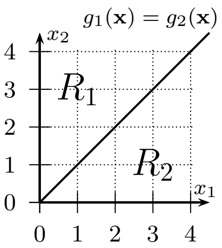
6 $\boxed{\text{C}}$ En la figura se representan frontera y regiones de decisión de un clasificador binario. ¿Cuál de los siguientes pares de vectores de pesos corresponde al clasificador de la figura?

A) $\mathbf{w}_1 = (-1, -1, -2)^t$ y $\mathbf{w}_2 = (-1, -2, -1)^t$ $\quad x_2 = x_1 \quad R_1 : x_2 < x_1 \quad R_2 : x_2 > x_1$

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

C) $\mathbf{w}_1 = (\ \ 1, \ \ 1, \ \ 2)^t$ y $\mathbf{w}_2 = (\ \ 1, \ \ 2, \ \ 1)^t$ $\quad x_2 = x_1 \quad R_1 : x_2 > x_1 \quad R_2 : x_2 < x_1$

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



7 $\boxed{\text{D}}$ Sea un problema de clasificación en 3 clases, $c = 1, 2, 3$, para objetos representados mediante vectores de características bidimensionales. Se tienen 3 muestras de entrenamiento representadas en notación homogénea: $\mathbf{x}_1 = (1, 1, 2)^t$ de la clase $c_1 = 1$, $\mathbf{x}_2 = (1, 2, 3)^t$ de la clase $c_2 = 2$ y $\mathbf{x}_3 = (1, 3, 1)^t$ de la clase $c_3 = 3$. Asimismo, se tiene un clasificador lineal definido por los vectores de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (2, -8, 0)^t$, $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-5, -2, -1)^t$ y $\mathbf{w}_3 = (w_{30}, w_{31}, w_{32}) = (-2, 1, -10)^t$. Si aplicamos una iteración del algoritmo Perceptrón a partir de estos vectores de pesos, con factor de aprendizaje $\alpha = 1$ y margen $b = 1.5$, entonces:

A) Se modificarán los vectores de pesos $\mathbf{w}_1$ y $\mathbf{w}_2$.

B) Se modificarán los vectores de pesos $\mathbf{w}_1$ y $\mathbf{w}_3$.

C) Se modificarán los vectores de pesos $\mathbf{w}_2$ y $\mathbf{w}_3$.

D) No se modificará ningún vector de pesos.

8 D En el proceso de entrenamiento de un árbol de clasificación, un nodo interno t tiene un grado de impureza $\mathcal{I}(t) > 0$. Uno de los "splits" produce un decremento de impureza igual a $\mathcal{I}(t)$. Indica la afirmación correcta:

A) No es posible lograr ese decremento de impureza.
B) Dicho "split" genera dos nodos impuros.
C) Dicho "split" genera un nodo puro y otro impuro.
D) Dicho "split" genera dos nodos puros.

9 A Para un problema de clasificación de datos bidimensionales $\mathbf{x} = (x_1, x_2)$ en dos clases disponemos de un árbol de clasificación. ¿Qué tipo de fronteras de decisión define el nodo raíz?

A) $a \cdot x_1 + b \cdot x_2 + c = 0$ donde $a = 0 \ \vee \ b = 0$ Fronteras paralelas a los ejes
B) $a \cdot x_1 + b \cdot x_2 + c = 0$ donde $a \neq 0 \ \wedge \ b \neq 0$
C) $a \cdot x_1 + b \cdot x_2 + c = 0$ donde $a \neq 0 \ \vee \ b = 0$
D) $a \cdot x_1 + b \cdot x_2 + c = 0$ donde $a = 0 \ \vee \ b \neq 0$

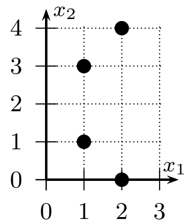
10 D Supóngase que estamos aplicando el algoritmo de aprendizaje de árboles de clasificación para un problema de cuatro clases: 1, 2, 3 y 4. El algoritmo ha alcanzado un nodo t que incluye un dato de cada clase, esto es, 4 en total. Se pretende evaluar la calidad de una partición del nodo t mediante un *"split"* $s = (j, r)$, que divide los datos en dos nodos t_1 y t_2 de la siguiente forma: los datos de las clases 1 y 2 quedan en el nodo t_1 y los datos de las clases 3 y 4 quedan en el nodo t_2. El decremento de impureza $\Delta\mathcal{I}(j, r, t)$ (medida como entropía) para cuantificar la calidad de esta partición es:

A) $\Delta\mathcal{I}(j, r, t) < 0.0$.
B) $0.0 \leq \Delta\mathcal{I}(j, r, t) < 0.5$.
C) $0.5 \leq \Delta\mathcal{I}(j, r, t) < 1.0$.
D) $1.0 \leq \Delta\mathcal{I}(j, r, t)$. $\Delta\mathcal{I}(j, r, t) = \mathcal{I}(t) - \hat{P}_t(t_1)\mathcal{I}(t_1) - \hat{P}_t(t_2)\mathcal{I}(t_2) = 2 - \frac{1}{2} \cdot 1 - \frac{1}{2} \cdot 1 = 1$

11 A Indica cuaĺ de las siguientes afirmaciones sobre un árbol de clasificación construido mediante el algoritmo de aprendizaje de árboles es *incorrecta*.

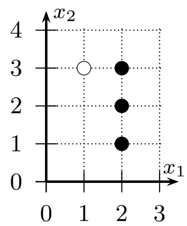
A) En cada nodo t, la probabilidad a posteriori de cualquier clase c, $P(c \mid t)$, es siempre mayor o igual que el menor de los pesos o probabilidades de decisión de sus dos hijos.
B) En cada nodo t la suma para todas las clases de $P(c \mid t)$ es 1.
C) La impureza de un nodo, medida como entropía, no puede ser menor que 0 ni mayor que $\log_2 C$, donde C es el número de clases.
D) Si N es el número de datos de aprendizaje, la profundidad del árbol no será mayor que N aunque, en la práctica, suele ser proporcional a $\log_2 N$.

12 D En la figura de la derecha se representan 4 muestras de bidimensionales. ¿Cuál es el número de clústers que minimiza la suma de errores cuadráticos para dicho conjunto de muestras?



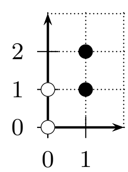
A) 1
B) 2
C) 3
D) 4 $J = 0$

13 D La figura a la derecha muestra una partición de 4 puntos bidimensionales en 2 clústers (representados mediante los símbolos • y ∘). La transferencia del punto $(2, 3)^t$ del clúster • al ∘ conduce a una variación de la SEC, ΔJ, tal que:



A) $\Delta J > 0$.
B) $0 \geq \Delta J > -\frac{1}{2}$.
C) $-\frac{1}{2} \geq \Delta J > -1$.
D) $-1 \geq \Delta J$. $\Delta J = \frac{1}{2} - \frac{3}{2} = -1$

14 A En la figura de la derecha se muestra una partición de 4 puntos bidimensionales de 2 clústers. La transferencia del punto $(1, 1)^t$ del clúster • al clúster ∘



A) produce un incremento en la SEC.
B) produce un decremento en la SEC.
C) no altera la SEC.
D) produce una SEC negativa.

15 A Considérese el algoritmo C-medias de Duda y Hart. Indicar cuál de las siguientes afirmaciones es *correcta*:

A) Su buena eficacia computational se consigue gracias al cálculo incremental de la variación de distorsión y de los vectores media de clúster.
B) Determina el número de clústers que minimiza la suma de errores cuadráticos (SEC).
C) Cuando un clúster se queda vacío, dicho clúster se elimina.
D) Ninguna de las anteriores.