

Examen del bloque 2 de SIN: Test (1,75 puntos)

ETSINF, Universitat Politècnica de València, 18 de enero de 2021

Grupo, apellidos y nombre: 3B, Arnal García, David

Marca cada recuadro con una única opción. Puntuación: $\max(0, (\text{aciertos} - \text{errores} / 3) \cdot 1,75 / 9)$.

- 1 ☐ Sea un problema de clasificación en tres clases para datos del tipo $\mathbf{x} = (x_1, x_2)^t \in \{0, 1\}^2$, con las distribuciones de probabilidad de la tabla. Indica en qué intervalo se halla el error de Bayes, ε^* :

\mathbf{x}		$P(c \mathbf{x})$			$P(\mathbf{x})$
x_1	x_2	$c=1$	$c=2$	$c=3$	
0	0	0.4	0.5	0.1	0.4
0	1	0.4	0.1	0.5	0.1
1	0	0.1	0.6	0.3	0.1
1	1	0.5	0.1	0.4	0.4

A) $\varepsilon^* < 0.40$.

B) $0.40 \leq \varepsilon^* < 0.45$.

C) $0.45 \leq \varepsilon^* < 0.50$.

D) $0.50 \leq \varepsilon^*$.

- 2 ☐ Supóngase que estamos aplicando el algoritmo de aprendizaje de árboles de clasificación para un problema de cuatro clases, $c = 1, 2, 3, 4$. El algoritmo ha alcanzado un nodo t que incluye los siguientes datos: 2 de la clase 1, 4 de la 2, 32 de la 3 y 256 de la 4. La impureza de t , $\mathcal{I}(t)$, medida como la entropía de la distribución empírica de las probabilidades a posteriori de las clases en t , es:

A) $0.00 \leq \mathcal{I}(t) < 0.50$.

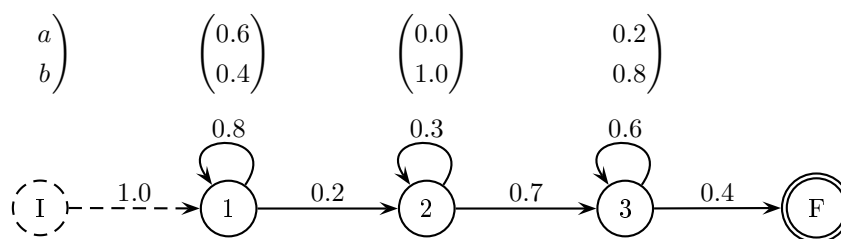
B) $0.50 \leq \mathcal{I}(t) < 1.00$.

C) $1.00 \leq \mathcal{I}(t) < 1.50$.

D) $1.50 \leq \mathcal{I}(t)$.

2

- 3 ☐ Sean M un modelo de Markov de representación gráfica:



¿Cuántas cadenas distintas de longitud 3 puede generar M ?

A) Ninguna.

a-b-a

B) Al menos una, pero no más de 3.

a-b-b

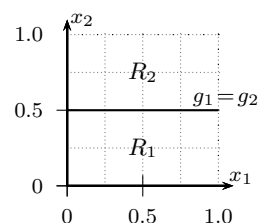
C) Más de 3, pero no más de 6.

b-b-a

D) Más de 6.

b-b-b

- 4 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos (en notación homogénea) define un clasificador equivalente al dado?



- A) $\mathbf{w}_1 = (1, 0, 0)^t$ y $\mathbf{w}_2 = (0, 0, 2)^t$.
 B) $\mathbf{w}_1 = (-1, 0, 0)^t$ y $\mathbf{w}_2 = (0, 0, -2)^t$.
 C) $\mathbf{w}_1 = (0, 0, 2)^t$ y $\mathbf{w}_2 = (1, 0, 0)^t$.
 D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

- 5 ☐ Sea \mathbf{x} un objeto a clasificar en una clase de C posibles. Indica cuál de los siguientes clasificadores *no* es de error mínimo (o escoge la última opción si los tres son de error mínimo):

- A) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(c) p(\mathbf{x}|c)$
 B) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} \log p(\mathbf{x}, c)$
 C) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(\mathbf{x}|c)$
 D) Los tres clasificadores anteriores son de error mínimo.

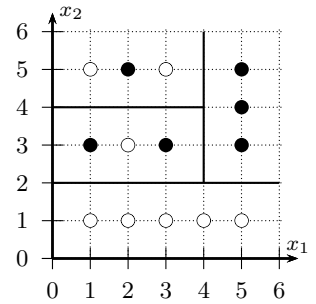
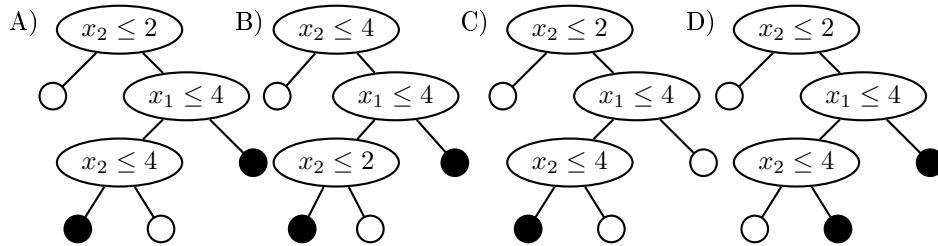
- 6 ☐ Supóngase que tenemos dos cajas con 50 manzanas cada una. La primera caja contiene 19 manzanas Gala y 31 Fuji. La segunda caja contiene 25 manzanas de cada tipo. Ahora supóngase que se escoge una caja al azar, y luego una manzana al azar de la caja escogida. Si la manzana escogida es Gala, la probabilidad P de que proceda de la primera caja es:

- A) $0/4 \leq P < 1/4$.
 B) $1/4 \leq P < 2/4$.
 C) $2/4 \leq P < 3/4$.
 D) $3/4 \leq P \leq 4/4$.

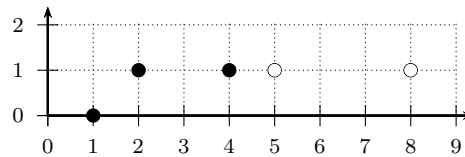
- 7 ☐ La probabilidad de error de un clasificador se estima que es del 12 %. Determina cuál es el número mínimo de muestras de test necesario, M , para conseguir que el intervalo de confianza al 95 % de dicho error no supere el ± 1 %; esto es, $I = [11 \%, 13 \%$:

- A) $M < 2000$.
 B) $2000 \leq M < 3500$.
 C) $3500 \leq M < 5000$.
 D) $M \geq 5000$.

- 8 ☐ Dado el conjunto de muestras de 2 clases (\circ y \bullet) de la figura de la derecha, ¿cuál de los siguientes árboles de clasificación es coherente con la partición representada?



- 9 ☐ La figura siguiente muestra una partición de 5 puntos bidimensionales de dos clústers, \bullet y \circ :



La transferencia del punto $(4,1)^t$ del clúster \bullet al clúster \circ produce una variación de la suma de errores cuadráticos, ΔJ , tal que:

- A) $\Delta J < 0$, esto es, la transferencia es provechosa.
 B) $0 \leq \Delta J < 1$.
 C) $1 \leq \Delta J < 2$.
 D) $\Delta J \geq 2$.

Examen del bloque 2 de SIN: Problema (2 puntos)

ETSINF, Universitat Politècnica de València, 18 de enero de 2021

Grupo, apellidos y nombre: 3B, Arnal García, David

Problema sobre Perceptrón

En la tabla siguiente se proporciona un conjunto de 4 muestras bidimensionales de aprendizaje de 3 clases, $c = 1, 2, 3$.

n	x_{n1}	x_{n2}	c_n
1	1	3	1
2	1	2	2
3	5	2	3
4	2	5	2

Se pide:

- (1.5 puntos) Realiza una traza de ejecución del algoritmo Perceptrón, hasta 3 iteraciones, con factor de aprendizaje $\alpha = 1$, margen $\gamma = 0.1$ y pesos iniciales nulos.
- (0.5 puntos) Clasifica la muestra de test $\mathbf{x} = (2, 1)^t$ mediante un clasificador lineal con los vectores de pesos obtenidos tras la tercera iteración.