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

ETSINF, Universitat Politècnica de València, 19 de diciembre de 2024

Grupo, apellidos y nombre: 2,

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

1 $\boxed{\text{B}}$ Dada la siguiente tabla de probabilidades condicionales de las 3 variables de interés:

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
P(A, B | C)	0.125	0.188	0.375	0.312	0.408	0.190	0.092	0.310

Si $P(C=0) = 0.72$, ¿cuál es el valor de $P(A=0 \mid B=0, C=1)$? $P(A=0 \mid B=0, C=1) = 0.497$

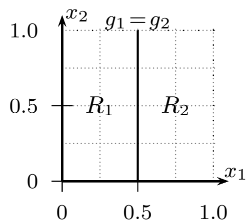
A) $\text{P(A=0} \mid B=0, C=1) \leq 0.25$

B) $0.25 < \text{P(A=0} \mid B=0, C=1) \leq 0.50$

C) $0.50 < \text{P(A=0} \mid B=0, C=1) \leq 0.75$

D) $0.75 < \text{P(A=0} \mid B=0, C=1) \leq 1.00$

2 $\boxed{\text{B}}$ 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) $\mathbf{w}_1 = (-0.5, 0, 0)^t$ y $\mathbf{w}_2 = (0, -1, 0)^t$.

B) $\mathbf{w}_1 = (0.5, 0, 0)^t$ y $\mathbf{w}_2 = (0, 1, 0)^t$.

C) $\mathbf{w}_1 = (0, 1, 0)^t$ y $\mathbf{w}_2 = (0.5, 0, 0)^t$.

D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

3 $\boxed{\text{B}}$ Sea un problema de clasificación en dos 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 la probabilidad de error ε del clasificador $c(\mathbf{x})$ basado en la función discriminante $g(\mathbf{x}) = 1.0 - x_1 + 0.5x_2$ definido como

$$c(\mathbf{x}) = \begin{cases} 1 & \text{si } g(\mathbf{x}) < 0 \\ 2 & \text{en caso contrario} \end{cases}$$

$\mathbf{x}$		$P(c \mid \mathbf{x})$		
x_1	x_2	$c=1$	$c=2$	$P(\mathbf{x})$
0	0	0.9	0.1	0
0	1	0.8	0.2	0.1
1	0	0.1	0.9	0.5
1	1	0.6	0.4	0.4

$\varepsilon = 0.37$

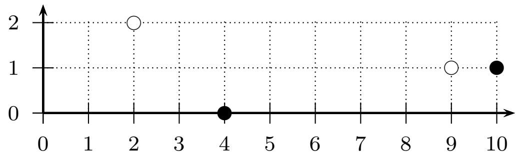
A) $\varepsilon < 0.25$.

B) $0.25 \leq \varepsilon < 0.50$.

C) $0.50 \leq \varepsilon < 0.75$.

D) $0.75 \leq \varepsilon$.

4 $\boxed{\text{B}}$ La figura siguiente muestra una partición de 4 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Si transferimos de clúster el punto $(10, 1)^t$, se produce una variación de la suma de errores cuadráticos (SEC), $\Delta J = J - J'$ (SEC tras el intercambio menos SEC antes del intercambio), tal que:

A) $\Delta J < -7$. $\qquad$ $\Delta J = 38.7 - 43.5 = -4.8$

B) $-7 \leq \Delta J < 0$.

C) $0 \leq \Delta J < 7$.

D) $\Delta J \geq 7$.

5 $\boxed{\text{D}}$ Sea $g(\mathbf{x})$ un clasificador. Indica cuál de las siguientes funciones *no* define un clasificador equivalente (o escoge la última opción si todas definen un clasificador equivalente):

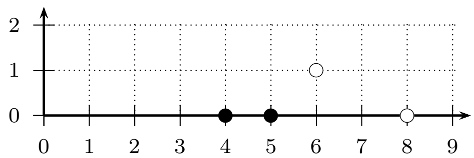
A) $f(g(\mathbf{x})) = ag(\mathbf{x}) + b \quad a > 0$

B) $f(g(\mathbf{x})) = a^{g(\mathbf{x})} \quad a > 1$

C) $f(g(\mathbf{x})) = ag(\mathbf{x})^3 \quad a > 0$

D) Las tres funciones anteriores definen un clasificador equivalente.

6 $\boxed{\text{A}}$ La figura siguiente muestra una partición de 4 puntos bidimensionales en dos clústers, $\bullet$ y $\circ$:



Indica cuál de los siguientes puntos se transfiere de clúster cuando aplicamos el algoritmo K-medias de Duda y Hart, pero no cuando aplicamos la versión convencional del algoritmo K-medias:

A) $(6, 1)^t$

B) $(4, 0)^t$

C) $(8, 0)^t$

D) $(5, 0)^t$

7 $\boxed{\text{B}}$ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0, 1, -2)^t$, $\mathbf{w}_2 = (0, -1, 2)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los mismos vectores de pesos, ¿cuál de las siguientes es esa última muestra?

A) $((5, 4)^t, 1)$
B) $((1, 1)^t, 2)$
C) $((2, 1)^t, 1)$
D) $((1, 4)^t, 1)$

8 $\boxed{\text{C}}$ Supóngase que tenemos una caja con 10 naranjas que contiene 4 naranjas Powell (P) y 6 Valencia (V) de la que extraemos dos naranjas, una detrás de otra sin reposición. Dadas las variables aleatorias:

- N1: variedad de la primera naranja extraída.
- N2: variedad de la segunda naranja extraída.

¿Cuál de las siguientes condiciones no es cierta?

A) $P(N2 = P) < P(N2 = P \mid N1 = V)$
B) $P(N1 = P,\, N2 = V) = P(N1 = V,\, N2 = P)$
C) $P(N1 = V) = P(N1 = V \mid N2 = P)$
D) $P(N2 = P) > P(N2 = P \mid N1 = P)$

9 $\boxed{\text{D}}$ 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 ninguno de los tres es de error mínimo):

A) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\min}\ e^{p(\mathbf{x},c)}$
B) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\max}\ -\log p(\mathbf{x}, c)$
C) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\min}\ \log p(\mathbf{x}, c)$
D) Ninguno de los tres clasificadores anteriores es de error mínimo.

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

ETSINF, Universitat Politècnica de València, 19 de diciembre de 2024

Grupo, apellidos y nombre: 2,

Problema sobre regresión logística

La siguiente tabla presenta por fila una muestra de entrenamiento de 2 dimensiones procedente de una clase:

n	x_{n1}	x_{n2}	c_n
1	1	1	2

Adicionalmente, la siguiente tabla representa una matriz de pesos iniciales con los pesos de cada clase dispuestos por columnas:

$\mathbf{w}_1$	$\mathbf{w}_2$
-0.5	0.5
-0.5	0.5
-0.5	0.5

Se pide:

1. (0.25 puntos) Calcula el vector de logits asociado a la muestra de entrenamiento.
2. (0.25 puntos) Aplica la función softmax al vector de logits de la muestra de entrenamiento.
3. (0.25 puntos) Calcula la neg-log-verosimilitud del conjunto de entrenamiento respecto a la matriz de pesos iniciales.
4. (0.25 puntos) Clasifica la muestra de entrenamiento. En caso de empate, elige cualquier clase.
5. (0.5 puntos) Calcula el gradiente de la función NLL en el punto de la matriz de pesos iniciales.
6. (0.5 puntos) Actualiza la matriz de pesos iniciales aplicando descenso por gradiente con factor de aprendizaje $\eta = 1.0$.

Solución:

1. Vector de logits de la muestra de entrenamiento:

n	a_{n1}	a_{n2}
1	-1.5	1.5

2. Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.05	0.95

3. Cálculo de la neg-log-verosimilitud:

$$\text{NLL}(\mathbf{W}) = 0.05$$

4. Clasificación de la muestra de entrenamiento:

n	$\hat{c}(x_n)$
1	2

5. Gradiente:

$\mathbf{g}_1$	$\mathbf{g}_2$
0.05	-0.05
0.05	-0.05
0.05	-0.05

6. Matriz de pesos actualizada:

$\mathbf{w}_1$	$\mathbf{w}_2$
-0.55	0.55
-0.55	0.55
-0.55	0.55