

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: 1,

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

- 1 ☒ 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 $\mathbf{w}_1 = (0, 0, -2)^t$, $\mathbf{w}_2 = (0, 0, 2)^t$. 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) $((5, 5)^t, 1)$
B) $((2, 4)^t, 1)$
C) $((2, 5)^t, 2)$
D) $((4, 1)^t, 1)$

- 2 ☐ 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)	0.449	0.173	0.051	0.327	0.343	0.027	0.157	0.473

Si $P(C = 0) = 0.81$, ¿cuál es el valor de $P(A = 1 | B = 0, C = 1)$? $P(A = 1 | B = 0, C = 1) = 0.135$

- A) $P(A=1 | B = 0, C = 1) \leq 0.25$
B) $0.25 < P(A=1 | B = 0, C = 1) \leq 0.50$
C) $0.50 < P(A=1 | B = 0, C = 1) \leq 0.75$
D) $0.75 < P(A=1 | B = 0, C = 1) \leq 1.00$

- 3 ☒ 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 $g(\mathbf{x}) = 0.5 + x_1 + x_2$ definido como

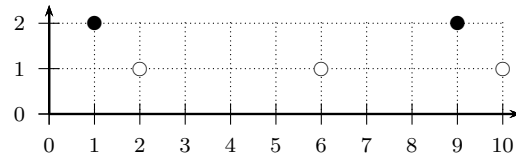
$$c(\mathbf{x}) = \begin{cases} 1 & \text{si } g(\mathbf{x}) < 0 \\ 2 & \text{en caso contrario} \end{cases}$$

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

$\varepsilon = 0.65$

- A) $\varepsilon < 0.25$.
B) $0.25 \leq \varepsilon < 0.50$.
C) $0.50 \leq \varepsilon < 0.75$.
D) $0.75 \leq \varepsilon$.

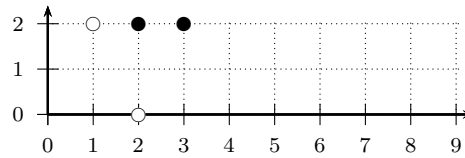
- 4 **A** La figura siguiente muestra una partición de 5 puntos bidimensionales en dos clústers, \bullet y \circ :



Si transferimos de clúster el punto $(9, 2)^t$, 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:

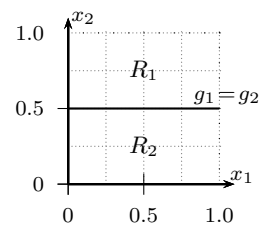
- A) $\Delta J < -7$. $\Delta J = 39.5 - 64.0 = -24.5$
 B) $-7 \leq \Delta J < 0$.
 C) $0 \leq \Delta J < 7$.
 D) $\Delta J \geq 7$.

- 5 **D** 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) $(2, 0)^t$
 B) $(2, 2)^t$
 C) $(3, 2)^t$
 D) $(1, 2)^t$
- 6 **C** 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) $\mathbf{w}_1 = (0, 0, -2)^t$ y $\mathbf{w}_2 = (-1, 0, 0)^t$.
 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.

7 ☒ C Supóngase que tenemos una caja con 10 naranjas que contiene 8 naranjas Washington (W) y 2 Cadenera (C) 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) $P(N1 = W, N2 = C) = P(N1 = C, N2 = W)$
- B) $P(N2 = W) < P(N2 = W \mid N1 = C)$
- C) $P(N1 = C) = P(N1 = C \mid N2 = W)$
- D) $P(N2 = W) > P(N2 = W \mid N1 = W)$

8 ☒ 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) $c(\mathbf{x}) = \arg \min_{c=1, \dots, C} e^{p(c|\mathbf{x})} + e^{p(\mathbf{x})}$
- B) $c(\mathbf{x}) = \arg \min_{c=1, \dots, C} e^{p(\mathbf{x}, c)}$
- C) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} -\log p(\mathbf{x}, c)$
- D) Ninguno de los tres clasificadores anteriores es de error mínimo.

9 ☒ B, 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) $f(g(\mathbf{x})) = \log g(\mathbf{x}) \quad g(\mathbf{x}) > 0$
- C) $f(g(\mathbf{x})) = \exp g(\mathbf{x})$
- D) Las tres funciones anteriores definen un clasificador equivalente.

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

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

Grupo, apellidos y nombre: 1,

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	1

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:

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

Solución:

- Vector de logits de la muestra de entrenamiento:

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

- Aplicación de la función softmax:

n	μ_{n1}	μ_{n2}
1	0.95	0.05

- Cálculo de la neg-log-verosimilitud:

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

- Clasificación de la muestra de entrenamiento:

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

- Gradiente:

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

- Matriz de pesos actualizada:

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