Examen de Teoría de Percepción - Segundo Parcial

ETSINF, Universitat Politècnica de València, Junio de 2022

Apellidos: ______ **Nombre:** ______

Profesor: ☐ **Jorge Civera** ☐ **Carlos Martínez**

Cuestiones (1.5 puntos, 30 minutos, sin apuntes)

D Se tiene el siguiente conjunto de muestras extraídas de una distribución de Bernoulli:

$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$
0	0	0	1	1	0	0	0
1	1	0	0	0	1	0	1

¿Cuál es el valor del parámetro de Bernoulli estimado por máxima verosimilitud?

A) $\mathbf{p} = \left(\frac{5}{8}\ \frac{3}{8}\right)$
B) $\mathbf{p} = \left(\frac{1}{2}\ \frac{1}{2}\ 0\ \frac{1}{2}\ \frac{1}{2}\ \frac{1}{2}\ 0\ \frac{1}{2}\right)$
C) $\mathbf{p} = \left(\frac{3}{4}\ \frac{1}{2}\right)$
D) $\mathbf{p} = \left(\frac{1}{4}\ \frac{1}{2}\right)$

C Si estamos suavizando una distribución Bernoulli con parámetro $\hat{\mathbf{p}} = (0.1\ 0.9\ 0.5)$ mediante truncamiento simple, ¿qué valor del hiperparámetro ϵ se ha utilizado si $\tilde{\mathbf{p}} = (0.2\ 0.8\ 0.5)$?

A) $\epsilon = 0.05$
B) $\epsilon = 0.10$
C) $\epsilon = 0.20$
D) $\epsilon = 0.30$

B ¿Qué finalidad tiene el coeficiente multinomial $\binom{x+}{x}$?

A) Convertir secuencias a vectores de contadores
B) Ponderar la probabilidad de un vector de contadores de acuerdo al número de secuencias diferentes que se pueden generar con ese vector de contadores
C) Aumentar la probabilidad de aquellos vectores de contadores cuyo número de secuencia diferentes asociado es menor
D) Suavizar la distribución multinomial para evitar probabilidades ceros de aquellos vectores de contadores que no se observan en el conjunto de entrenamiento.

$\boxed{\text{A}}$ Dado un problema de clasificación en un espacio bidimensional entre dos clases A y B equiprobables, con probabilidades condicionadas gaussianas de parámetros $\mu_A = (0\ -1)$ y $\mu_B = (1\ 0)$, y $\Sigma = I$ común a ambas clases, la frontera de decisión entre las dos regiones definidas por las distribuciones viene dada por:

A) $x_2 = -x_1$
B) $x_2 = \frac{1}{2}x_1^2 - x_1$
C) $x_2 = x_1$
D) $x_2 = 1 - x_1$

$\boxed{\text{C}}$ En un suavizado de la matriz de covarianzas por *flat smoothing* se interpola la matriz de covarianzas estimada para cada clase con:

A) La matriz identidad
B) La matriz de covarianza global
C) La matriz de covarianza global suavizada
D) La matriz de covarianza estimada de la clase de media más próxima

$\boxed{\text{D}}$ Dadas dos funciones Kernel K_1 y K_2, indicar cuál de las siguientes **no** es una función Kernel:

A) $3K_1(\mathbf{x}, \mathbf{y}) + \exp(K_2(\mathbf{x}, \mathbf{y}))$
B) $K_1(\mathbf{x}, \mathbf{y})^2 + K_2(\mathbf{x}, \mathbf{y})^4$
C) $\mathbf{x}^t\mathbf{x}K_1(\mathbf{x}, \mathbf{y}) \cdot K_2(\mathbf{x}, \mathbf{y})\mathbf{y}^t\mathbf{y}$
D) $K_2(\mathbf{x}, \mathbf{y})^3 - 5K_1(\mathbf{x}, \mathbf{y})^5$

$\boxed{\text{A}}$ Sabiendo que $S_b = \sum_{c=1}^{C} N_c\,(\overline{\mathbf{x}}_c - \overline{\mathbf{x}})(\overline{\mathbf{x}}_c - \overline{\mathbf{x}})^t$ y $S_w = \sum_{c=1}^{C} \Sigma_c$, ¿en qué consiste el criterio de optimización de LDA?

A) Maximizar S_b y minimizar S_w
B) Maximizar S_b y maximizar S_w
C) Minimizar S_b y minimizar S_w
D) Minimizar S_b y maximizar S_w

$\boxed{\text{B}}$ Los clasificadores débiles:

A) Presentan gran flexibilidad
B) Tienen un número de parámetros relativamente bajo
C) Son, entre otros, clasificadores como el de k vecinos más cercanos
D) Sólo son aplicables a datos vectoriales

Examen de Teoría de Percepción - Segundo Parcial

ETSINF, Universitat Politècnica de València, Junio de 2022

Apellidos: ______________________ **Nombre:** ____________

Profesor: ☐ **Jorge Civera** ☐ **Carlos Martínez**

Problemas (2 puntos, 90 minutos, con apuntes)

1. (**1 punto**) Se tiene el siguiente conjunto de imágenes de 4×4 píxeles representado a 8 niveles de gris pertenecientes a imágenes de palos de la baraja francesa (tréboles, diamantes, corazones y picas):

	7	5	6	7	7	2	2	7	7	3	5	7	7	6	6	7	7	6	6	7
Tréboles	6	2	3	7	5	0	1	5	6	3	4	6	4	2	2	4	5	3	3	5
	6	3	3	6	2	0	0	2	5	3	4	6	6	3	2	6	4	3	3	4
	7	7	7	7	5	3	3	6	6	5	5	7	7	5	4	7	7	7	7	7
	7	7	7	7	7	5	5	7	7	5	5	7	7	6	5	7	7	7	7	7
Diamantes	6	5	5	6	6	2	2	6	6	2	2	5	6	3	2	5	6	4	4	5
	6	5	5	6	6	3	3	6	6	2	2	6	5	2	2	6	6	5	5	6
	7	7	7	7	7	6	6	7	7	5	5	7	7	5	6	7	7	7	7	7
	7	6	6	7	4	3	3	4	7	6	6	7	3	3	3	3	6	6	6	6
Corazones	6	2	2	6	3	2	2	3	6	2	3	6	2	2	2	3	5	4	4	5
	7	4	4	7	6	2	2	6	5	1	1	5	4	3	4	5	6	5	5	6
	7	7	7	7	7	5	5	7	5	4	3	6	7	6	6	6	7	7	7	7
	7	6	6	7	7	4	4	7	7	6	6	7	7	6	6	7	7	6	6	7
Picas	6	1	1	6	4	0	0	5	6	2	2	6	3	3	3	3	5	3	3	5
	5	1	1	5	2	0	0	3	5	2	2	4	4	3	3	4	4	2	2	4
	7	7	7	7	6	3	3	6	7	5	5	7	7	6	6	7	7	6	6	7

Se pide:

a) Realizar la representación por histograma de niveles de gris de las distintas imágenes (**0.2 puntos**)

b) Estimar los parámetros de las distribuciones multinomiales de cada clase tomando como datos de entrenamiento las cuatro primeras imágenes presentadas de cada una de ellas (**0.5 puntos**)

c) Calcular la tasa de error del clasificador estimado al clasificar la última imagen de cada clase, considerando las clases equiprobables. Asume $0^0 = 0$ y $0 \log(0) = 0$. (**0.3 puntos**)

Solución:

a)

Tréboles								Diamantes								Corazones								Picas							
0	0	1	3	0	1	4	7	0	0	0	0	0	4	4	8	0	0	2	0	2	0	4	8	0	4	0	0	0	2	4	6
3	1	4	2	0	3	1	2	0	0	2	2	0	2	6	4	0	0	4	4	2	2	2	2	4	0	1	3	3	1	2	2
0	0	0	3	2	4	4	3	0	0	4	0	0	5	3	4	0	2	1	2	1	3	5	2	0	0	4	0	1	3	4	4
0	0	3	1	3	1	4	4	0	0	3	1	0	4	4	4	0	0	3	6	2	1	3	1	0	0	0	6	2	0	4	4
0	0	0	4	2	2	2	6	0	0	0	0	2	3	3	8	0	0	0	0	2	4	6	4	0	0	2	2	2	2	4	4

b) $p_T = \left(\frac{3}{64}\ \frac{1}{64}\ \frac{8}{64}\ \frac{9}{64}\ \frac{5}{64}\ \frac{9}{64}\ \frac{13}{64}\ \frac{16}{64}\right)$
$p_D = \left(0\ 0\ \frac{9}{64}\ \frac{3}{64}\ 0\ \frac{15}{64}\ \frac{17}{64}\ \frac{20}{64}\right)$
$p_C = \left(0\ \frac{2}{64}\ \frac{10}{64}\ \frac{12}{64}\ \frac{7}{64}\ \frac{6}{64}\ \frac{14}{64}\ \frac{13}{64}\right)$
$p_P = \left(\frac{4}{64}\ \frac{4}{64}\ \frac{5}{64}\ \frac{9}{64}\ \frac{6}{64}\ \frac{6}{64}\ \frac{14}{64}\ \frac{16}{64}\right)$

c) Como las $P(c)$ son equiprobables, se puede usar $g_c(\mathbf{x}) = \prod_{d=1}^{D} p_{cd}^{x_d}$
Para trébol: $\mathbf{x} = (0\,0\,0\,4\,2\,2\,2\,6)^t$

- $g_T(\mathbf{x}) = 1 \cdot 1 \cdot 1 \cdot \frac{9^4}{64^4} \cdot \frac{5^2}{64^2} \cdot \frac{9^2}{64^2} \cdot \frac{13^2}{64^2} \cdot \frac{16^6}{64^6} = 5^2 \cdot 9^2 \cdot 13^2 \cdot \frac{9^4 \cdot 16^4}{64^{16}} = 342225 \cdot \frac{9^4 \cdot 16^4}{64^{16}}$
- $g_D(\mathbf{x}) = 0$
- $g_C(\mathbf{x}) = 0$
- $g_P(\mathbf{x}) = 1 \cdot 1 \cdot 1 \cdot \frac{9^4}{64^4} \cdot \frac{6^2}{64^2} \cdot \frac{6^2}{64^2} \cdot \frac{14^2}{64^2} \cdot \frac{16^6}{64^6} = 6^2 \cdot 6^2 \cdot 14^2 \cdot \frac{9^4 \cdot 16^4}{64^{16}} = 254016 \cdot \frac{9^4 \cdot 16^4}{64^{16}}$

Se clasifica como trébol, acierta.
Para diamantes y corazón, $g_D(\mathbf{x}) = g_C(\mathbf{x}) = 0$, con lo que falla para ambas.
Para pica: $\mathbf{x} = (0\,0\,2\,2\,2\,2\,4\,4)^t$

- $g_T(\mathbf{x}) = 1 \cdot 1 \cdot \frac{8^2}{64^2} \cdot \frac{9^2}{64^2} \cdot \frac{5^2}{64^2} \cdot \frac{9^2}{64^2} \cdot \frac{13^4}{64^4} \cdot \frac{16^4}{64^4} = 8^2 \cdot 5^2 \cdot 9^2 \cdot 13^4 \cdot \frac{9^2 \cdot 16^4}{64^{16}} = 3701505600 \cdot \frac{9^2 \cdot 16^4}{64^{16}}$
- $g_D(\mathbf{x}) = 0$
- $g_C(\mathbf{x}) = 0$
- $g_P(\mathbf{x}) = 1 \cdot 1 \cdot \frac{5^2}{64^2} \cdot \frac{9^2}{64^2} \cdot \frac{6^2}{64^2} \cdot \frac{6^2}{64^2} \cdot \frac{14^4}{64^4} \cdot \frac{16^4}{64^4} = 5^2 \cdot 6^2 \cdot 6^2 \cdot 14^4 \cdot \frac{9^2 \cdot 16^4}{64^{16}} = 1244678400 \cdot \frac{9^2 \cdot 16^4}{64^{16}}$

Se clasifica como trébol, falla.
Por tanto, el clasificador presenta una tasa de error del 75 %.

2. (**0.5 puntos**) Se tiene la siguiente matriz Gramm asociado al conjunto de muestras $\{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, -1), (\mathbf{x}_4, -1)$:

$$\mathbb{K} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{4} & 1 \end{pmatrix}$$

Se pide realizar dos iteraciones del algoritmo Kernel perceptron partiendo del conjunto de pesos $\alpha = (0\ 0\ 0\ 0)$ e indicar el conjunto de pesos resultante.

Solución:

La fórmula general del clasificador es $g(\mathbf{x}) = \sum_{n=1}^{N} \alpha_n c_n K(\mathbf{x}_n, \mathbf{x}) + \alpha_n c_n$

Primera iteración:

- $n = 1$: $g(\mathbf{x}_1) = 0, c_1 g(\mathbf{x}_1) = 0 \leq 0 \to \alpha = (1\ 0\ 0\ 0)$
- $n = 2$: $g(\mathbf{x}_2) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_2) + \alpha_1 c_1 = \frac{1}{2} + 1 = \frac{3}{2}, c_2 g(\mathbf{x}_2) = \frac{3}{2} > 0 \to \alpha = (1\ 0\ 0\ 0)$
- $n = 3$: $g(\mathbf{x}_3) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_3) + \alpha_1 c_1 = \frac{1}{2} + 1 = \frac{3}{2}, c_3 g(\mathbf{x}_3) = -\frac{3}{2} \leq 0 \to \alpha = (1\ 0\ 1\ 0)$
- $n = 4$: $g(\mathbf{x}_4) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_4) + \alpha_1 c_1 + \alpha_3 c_3 K(\mathbf{x}_3, \mathbf{x}_4) + \alpha_3 c_3 = \frac{1}{4} + 1 - \frac{1}{4} - 1 = 0, c_4 g(\mathbf{x}_4) = 0 \leq 0 \to \alpha = (1\ 0\ 1\ 1)$

Segunda iteración:

- $n = 1$: $g(\mathbf{x}_1) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_1) + \alpha_1 c_1 + \alpha_3 c_3 K(\mathbf{x}_3, \mathbf{x}_1) + \alpha_3 c_3 + \alpha_4 c_4 K(\mathbf{x}_4, \mathbf{x}_1) + \alpha_4 c_4 = 1 + 1 - \frac{1}{2} - 1 - \frac{1}{4} - 1 = -\frac{3}{4}, c_1 g(\mathbf{x}_1) = -\frac{3}{4} \leq 0 \to \alpha = (2\ 0\ 1\ 1)$
- $n = 2$: $g(\mathbf{x}_2) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_2) + \alpha_1 c_1 + \alpha_3 c_3 K(\mathbf{x}_3, \mathbf{x}_2) + \alpha_3 c_3 + \alpha_4 c_4 K(\mathbf{x}_4, \mathbf{x}_2) + \alpha_4 c_4 = 1 + 2 - \frac{1}{3} - 1 - \frac{1}{3} - 1 = \frac{1}{3}, c_2 g(\mathbf{x}_2) = \frac{1}{3} > 0 \to \alpha = (2\ 0\ 1\ 1)$
- $n = 3$: $g(\mathbf{x}_3) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_3) + \alpha_1 c_1 + \alpha_3 c_3 K(\mathbf{x}_3, \mathbf{x}_3) + \alpha_3 c_3 + \alpha_4 c_4 K(\mathbf{x}_4, \mathbf{x}_3) + \alpha_4 c_4 = 1 + 2 - 1 - 1 - \frac{1}{4} - 1 = -\frac{1}{4}, c_3 g(\mathbf{x}_3) = \frac{1}{4} > 0 \to \alpha = (2\ 0\ 1\ 1)$
- $n = 4$: $g(\mathbf{x}_4) = \alpha_1 c_1 K(\mathbf{x}_1, \mathbf{x}_4) + \alpha_1 c_1 + \alpha_3 c_3 K(\mathbf{x}_3, \mathbf{x}_4) + \alpha_3 c_3 + \alpha_4 c_4 K(\mathbf{x}_4, \mathbf{x}_4) + \alpha_4 c_4 = \frac{1}{2} + 2 - \frac{1}{4} - 1 - 1 - 1 = -\frac{3}{4}, c_4 g(\mathbf{x}_4) = \frac{3}{4} > 0 \to \alpha = (2\ 0\ 1\ 1)$

3. (**0.5 puntos**) Se tiene el siguiente conjunto de datos y clasificadores lineales:

$$\mathbf{x}_1 = ((0,1,0), -1), \mathbf{x}_2 = ((1,0,1), +1), \mathbf{x}_3 = ((1,0,-1), -1), \mathbf{x}_4 = ((1,-1,1), +1)$$

$$g_1(\mathbf{z}) = \begin{cases} +1 & z_1 > 0 \\ -1 & z_1 \leq 0 \end{cases} \quad g_2(\mathbf{z}) = \begin{cases} +1 & z_2 > -1 \\ -1 & z_2 \leq -1 \end{cases} \quad g_3(\mathbf{z}) = \begin{cases} +1 & z_3 > 1 \\ -1 & z_3 \leq 1 \end{cases} \quad g_4(\mathbf{z}) = \begin{cases} +1 & z_1 + z_2 + z_3 \geq 0 \\ -1 & z_1 + z_2 + z_3 < 0 \end{cases}$$

Se pide realizar una primera iteración de AdaBoost sobre estos datos y clasificadores, indicando la tabla de acierto y fallo por clasificador, el clasificador escogido, el error en primera iteración (ϵ_1), el peso del clasificador escogido (α_1) y los pesos de las muestras en la siguiente iteración ($\mathbf{w}^{(2)}$).

Solución:

Tabla acierto/fallo:

	g_1	g_2	g_3	g_4
$\mathbf{x}_1$	✓	X	✓	X
$\mathbf{x}_2$	✓	✓	X	✓
$\mathbf{x}_3$	X	X	✓	X
$\mathbf{x}_4$	✓	X	X	✓

Vector de pesos de muestras inicial: $\mathbf{w}^{(1)} = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$

Error ponderado por clasificador: $g_1 : \frac{1}{4}$, $g_2 : \frac{3}{4}$, $g_3 : \frac{2}{4}$, $g_4 : \frac{2}{4}$

Por tanto, escogemos g_1

Error de clasificación: $\epsilon_1 = \frac{1}{4}$

Peso del clasificador: $\alpha_1 = \frac{1}{2} \log 3$

Pesos de las muestras en la siguiente iteración:

	$w_n^{(1)} \exp(-c_n \alpha_1 C_1(\mathbf{x}_n)$	$\mathbf{w}^{(2)}$
$\mathbf{x}_1$	$\frac{1}{4}\frac{1}{\sqrt{3}}$	$\frac{1}{6}$
$\mathbf{x}_2$	$\frac{1}{4}\frac{1}{\sqrt{3}}$	$\frac{1}{6}$
$\mathbf{x}_3$	$\frac{1}{4}\sqrt{3}$	$\frac{1}{2}$
$\mathbf{x}_4$	$\frac{1}{4}\frac{1}{\sqrt{3}}$	$\frac{1}{6}$
Suma	$\frac{3}{4}\frac{1}{\sqrt{3}} + \frac{1}{4}\sqrt{3} = \frac{\sqrt{3}}{2}$	