

Examen de Teoría de Percepción - Segundo Parcial

ETSINF, Universitat Politècnica de València, 5 de junio de 2020

Apellidos: Díaz-Alejo León Nombre: Stéphane

Profesor: ☐ Jorge Civera ☐ Carlos Martínez

Problemas (4 puntos, 90 minutos, con apuntes)

1. (2 puntos) Se dispone de 12 secuencias de 5 tiradas de dados personalizados de 6 caras por parte de 3 jugadores (X, Y, Z):

Secuencia	1	2	3	4	5	6	7	8	9	10	11	12
1ª tirada	5	6	2	5	1	2	1	5	2	2	5	1
2ª tirada	6	3	3	5	6	2	2	2	3	4	5	2
3ª tirada	1	5	2	1	2	4	2	1	1	4	2	2
4ª tirada	1	2	1	6	5	1	1	1	3	5	2	5
5ª tirada	3	3	2	6	4	3	6	3	3	4	3	4
Jugador	X	X	X	X	Y	Y	Y	Y	Z	Z	Z	Z

Así pues:

- Calcular los parámetros (probabilidades *a priori* y prototipos multinomiales \mathbf{p}_c) del clasificador multinomial por estimación de máxima verosimilitud sobre las tiradas de cada jugador. (0.75 puntos)
- Suaviza las probabilidades de los prototipos multinomiales utilizando descuento absoluto ($\epsilon = 0.05$) mediante interpolación, usando como distribución generalizada la distribución multinomial incondicional (independiente de la clase) estimada a partir de las muestras de entrenamiento (0.5 puntos)
- Determinar qué jugador realizó la secuencia de tiradas “2 4 1 3 5” por máxima probabilidad a posteriori (0.5 puntos)
- ¿Era necesario aplicar algún tipo de suavizado? Razona la respuesta. (0.25 puntos)

Solución:

- Primero es necesario realizar la conversión de secuencia de 5 tiradas a vector de contadores sobre el conjunto de posibles valores de un dado de 6 caras:

Secuencia	1	2	3	4	5	6	7	8	9	10	11	12
Contador 1	2	0	1	1	1	1	2	2	1	0	0	1
Contador 2	0	1	3	0	1	2	2	1	1	1	2	2
Contador 3	1	2	1	0	0	1	0	1	3	0	1	0
Contador 4	0	0	0	0	1	1	0	0	0	3	0	1
Contador 5	1	1	0	2	1	0	0	1	0	1	2	1
Contador 6	1	1	0	2	1	0	1	0	0	0	0	0
Jugador	X	X	X	X	Y	Y	Y	Y	Z	Z	Z	Z

Al haber igualdad de muestras de entrenamiento de cada clase, tendremos $P(X) = P(Y) = P(Z) = \frac{1}{3}$, y los prototipos multinomiales de cada clase serán:

	p_X	p_Y	p_Z
Contador 1	0.2	0.3	0.1
Contador 2	0.2	0.3	0.3
Contador 3	0.2	0.1	0.2
Contador 4	0.0	0.1	0.2
Contador 5	0.2	0.1	0.2
Contador 6	0.2	0.1	0.0

- Los prototipos multinomiales de cada clase tras aplicar el descuento serán:

	p_X	p_Y	p_Z
Contador 1	0.15	0.25	0.05
Contador 2	0.15	0.25	0.25
Contador 3	0.15	0.05	0.15
Contador 4	0.00	0.05	0.15
Contador 5	0.15	0.05	0.15
Contador 6	0.15	0.05	0.00

y la masa de probabilidad descontada en cada clase es

m_X	m_Y	m_Z
0.25	0.30	0.25

El prototipo multinomial incondicional a la clase es

	p_g
Contador 1	$\frac{6}{30}$
Contador 2	$\frac{8}{30}$
Contador 3	$\frac{5}{30}$
Contador 4	$\frac{3}{30}$
Contador 5	$\frac{5}{30}$
Contador 6	$\frac{3}{30}$

Si distribuimos la masa de probabilidad descontada entre todos los parámetros de cada clase, $\tilde{p}_c = p_c + m_c \cdot p_g$

	\tilde{p}_X	\tilde{p}_Y	\tilde{p}_Z
Contador 1	$0.15 + 0.25 \cdot \frac{6}{30}$	$0.25 + 0.30 \cdot \frac{6}{30}$	$0.05 + 0.25 \cdot \frac{6}{30}$
Contador 2	$0.15 + 0.25 \cdot \frac{8}{30}$	$0.25 + 0.30 \cdot \frac{8}{30}$	$0.25 + 0.25 \cdot \frac{8}{30}$
Contador 3	$0.15 + 0.25 \cdot \frac{5}{30}$	$0.05 + 0.30 \cdot \frac{5}{30}$	$0.15 + 0.25 \cdot \frac{5}{30}$
Contador 4	$0.00 + 0.25 \cdot \frac{3}{30}$	$0.05 + 0.30 \cdot \frac{3}{30}$	$0.15 + 0.25 \cdot \frac{3}{30}$
Contador 5	$0.15 + 0.25 \cdot \frac{5}{30}$	$0.05 + 0.30 \cdot \frac{5}{30}$	$0.15 + 0.25 \cdot \frac{5}{30}$
Contador 6	$0.15 + 0.25 \cdot \frac{3}{30}$	$0.05 + 0.30 \cdot \frac{3}{30}$	$0.00 + 0.25 \cdot \frac{3}{30}$

- c) Para clasificar la tirada “2 4 1 3 5”, la convertimos en vector de contadores $\mathbf{y} = (111110)$ y aplicamos la regla del clasificador multinomial:

$$\hat{c}(\mathbf{y}) = \arg \max_c \mathbf{y}^t \log(\tilde{p}_c) + \log(P(c)) = Z$$

- d) Sí, era necesario suavizar los prototipos multinomiales porque había probabilidades cuyo valor era cero.

2. (1.25 puntos) Tenemos la muestra de entrenamiento $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, -1), (\mathbf{x}_4, -1)\}$, y la matriz de kernel:

$$\mathbf{K} = \begin{pmatrix} 1.0 & 0.1 & 0.1 & 0.3 \\ 0.1 & 1.0 & 0.5 & 0.1 \\ 0.1 & 0.5 & 1.0 & 0.1 \\ 0.3 & 0.1 & 0.1 & 1.0 \end{pmatrix}$$

Obtén el valor de α del algoritmo Kernel Perceptron tras cada muestra y en cada iteración hasta convergencia.

Solución:

Partimos de $\alpha = \{0, 0, 0, 0\}$, usando la función a optimizar $g(\mathbf{x}) = \sum_{i=1}^4 \alpha_i \cdot c_i \cdot K(\mathbf{x}, \mathbf{x}_i) + \alpha_i \cdot c_i$.

1ª iteración

- $g(\mathbf{x}_1) = 0 \rightarrow c_1 g(\mathbf{x}_1) \leq 0 \rightarrow \alpha_1 = 1 \rightarrow \alpha = \{1, 0, 0, 0\}$
- $g(\mathbf{x}_2) = \alpha_1 K(\mathbf{x}_2, \mathbf{x}_1) + \alpha_1 \rightarrow c_2 g(\mathbf{x}_2) > 0 \rightarrow \alpha = \{1, 0, 0, 0\}$
- $g(\mathbf{x}_3) = \alpha_1 K(\mathbf{x}_3, \mathbf{x}_1) + \alpha_1 \rightarrow c_3 g(\mathbf{x}_3) \leq 0 \rightarrow \alpha_3 = 1 \rightarrow \alpha = \{1, 0, 1, 0\}$
- $g(\mathbf{x}_4) = \alpha_1 K(\mathbf{x}_4, \mathbf{x}_1) - \alpha_3 K(\mathbf{x}_4, \mathbf{x}_3) + \alpha_1 - \alpha_3 \rightarrow c_4 g(\mathbf{x}_4) \leq 0 \rightarrow \alpha_4 = 1 \rightarrow \alpha = \{1, 0, 1, 1\}$

2ª iteración

- $g(\mathbf{x}_1) = \alpha_1 K(\mathbf{x}_1, \mathbf{x}_1) - \alpha_3 K(\mathbf{x}_1, \mathbf{x}_3) - \alpha_4 K(\mathbf{x}_1, \mathbf{x}_4) + \alpha_1 - \alpha_3 - \alpha_4 \rightarrow c_1 g(\mathbf{x}_1) \leq 0 \rightarrow \alpha_1 = 2 \rightarrow \alpha = \{2, 0, 1, 1\}$
- $g(\mathbf{x}_2) = \alpha_1 K(\mathbf{x}_2, \mathbf{x}_1) - \alpha_3 K(\mathbf{x}_2, \mathbf{x}_3) - \alpha_4 K(\mathbf{x}_2, \mathbf{x}_4) + \alpha_1 - \alpha_3 - \alpha_4 \rightarrow c_2 g(\mathbf{x}_2) \leq 0 \rightarrow \alpha_2 = 1 \rightarrow \alpha = \{2, 1, 1, 1\}$
- $g(\mathbf{x}_3) = \alpha_1 K(\mathbf{x}_3, \mathbf{x}_1) + \alpha_2 K(\mathbf{x}_3, \mathbf{x}_2) - \alpha_3 K(\mathbf{x}_3, \mathbf{x}_3) - \alpha_4 K(\mathbf{x}_3, \mathbf{x}_4) + \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 \rightarrow c_3 g(\mathbf{x}_3) \leq 0 \rightarrow \alpha_3 = 2 \rightarrow \alpha = \{2, 1, 2, 1\}$
- $g(\mathbf{x}_4) = \alpha_1 K(\mathbf{x}_4, \mathbf{x}_1) + \alpha_2 K(\mathbf{x}_4, \mathbf{x}_2) - \alpha_3 K(\mathbf{x}_4, \mathbf{x}_3) - \alpha_4 K(\mathbf{x}_4, \mathbf{x}_4) + \alpha_1 + \alpha_2 - \alpha_3 - \alpha_4 \rightarrow c_4 g(\mathbf{x}_4) > 0 \rightarrow \alpha = \{2, 1, 2, 1\}$

Es necesaria una tercera iteración, pero no hay cambios y se obtiene finalmente que $\alpha = \{2, 1, 2, 1\}$.

3. **(0.75 puntos)** Sea el siguiente conjunto de datos de \mathbb{R}^3 sobre las clases $\mathbb{C} = \{-1, +1\}$:

$$\mathcal{X} = \left\{ \mathbf{x}_1 = \left(\left(\frac{3}{2}, -\frac{1}{2}, \frac{3}{4} \right), +1 \right), \mathbf{x}_2 = \left(\left(-2, \frac{5}{4}, \frac{1}{2} \right), -1 \right), \mathbf{x}_3 = \left(\left(1, \frac{3}{2}, 0 \right), +1 \right), \mathbf{x}_4 = \left(\left(-1, -\frac{1}{2}, 0 \right), -1 \right) \right\}$$

Se dispone de los siguientes clasificadores:

$$g_1(\mathbf{x}) = \begin{cases} +1 & x_1 \cdot x_3 > 0 \\ -1 & \text{otro caso} \end{cases} \quad g_2(\mathbf{x}) = \begin{cases} +1 & x_1 + x_3 > -3 \\ -1 & \text{otro caso} \end{cases} \quad g_3(\mathbf{x}) = \begin{cases} +1 & x_1 \cdot x_2 > 0 \\ -1 & \text{otro caso} \end{cases} \quad g_4(\mathbf{x}) = \begin{cases} +1 & x_1 + x_2 > 1 \\ -1 & \text{otro caso} \end{cases}$$

Se pide:

- Mostrar la tabla acierto/error para los clasificadores y las muestras propuestas. **(0.25 puntos)**
- Aplicar la primera iteración del algoritmo AdaBoost, indicando clasificador C_1 , error asociado ϵ_1 y peso del clasificador α_1 . En caso de empate entre error de clasificadores, escoged el de menor índice. **(0.25 puntos)**
- Calcular los pesos de las muestras para la siguiente iteración $w^{(2)}$. **(0.25 puntos)**

Solución:

- a) Tabla acierto/error

	g_1	g_2	g_3	g_4
x_1	✓	✓	X	X
x_2	✓	X	✓	✓
x_3	X	✓	✓	✓
x_4	✓	X	X	✓

- b) Error por clasificador: $g_1 \rightarrow \frac{1}{4}$, $g_2 \rightarrow \frac{1}{2}$, $g_3 \rightarrow \frac{1}{2}$, $g_4 \rightarrow \frac{1}{4}$

Clasificador escogido: $C_1 = g_1$

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

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

- c) Pesos siguiente iteración: $w^{(2)} = \left(\frac{1}{6}, \frac{1}{6}, \frac{3}{6}, \frac{1}{6} \right)$