

Examen de Teoría de Percepción - Recuperación Segundo Parcial

ETSINF, Universitat Politècnica de València, Junio de 2019

Apellidos:

Nombre:

Profesor: ☐ Jorge Civera ☐ Carlos Martínez

Cuestiones (2 puntos, 30 minutos, sin apuntes)

☐ D Al aplicar funciones kernel:

- A) Se proyecta explícitamente a un espacio alternativo donde las muestras serán linealmente separables
- B) Se proyecta explícitamente a un espacio alternativo, aunque no se garantiza siempre si las muestras serán linealmente separables
- C) No se proyecta explícitamente a un espacio alternativo, aunque sí se garantiza la separabilidad lineal de las muestras en cualquier caso
- D) No se proyecta explícitamente a un espacio alternativo ni se garantiza la separabilidad lineal en todos los casos

☐ C Dada la función kernel $K(\mathbf{x}, \mathbf{y})$, con $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{D \times 1}$, indicar cuál de las siguientes funciones derivadas sería un kernel

- A) $1 + K(\mathbf{x}, \mathbf{y})^{-1}$
- B) $K(\mathbf{x}, \mathbf{y})^2 - K(\mathbf{x}, \mathbf{y})$
- C) $2\mathbf{x}^t K(\mathbf{x}, \mathbf{y}) \mathbf{y}$
- D) $-K(\mathbf{x}, \mathbf{y})$

☐ B En la estimación de un clasificador probabilístico, habitualmente se asume una distribución de probabilidad en uno de sus factores y se estiman sus parámetros. ¿Sobre qué término se suele hacer?:

- A) Sobre la probabilidad *a posteriori* $P(c|x)$
- B) Sobre la probabilidad condicionada $p(x|c)$
- C) Sobre la probabilidad *a priori* $P(c)$
- D) Sobre la probabilidad incondicional $p(x)$

☐ B ¿Cuál de las siguientes es una forma correcta del clasificador de Bernoulli de parámetros $\Theta = \{P(1), \dots, P(C), p_1, \dots, p_C\}$?

- A) $c^*(x) = \arg \max_{c=1, \dots, C} \prod_{d=1}^D p_{cd}^{x_d} (1 - p_{cd})^{(1-x_d)}$
- B) $c^*(x) = \arg \max_{c=1, \dots, C} P(c) \prod_{d=1}^D p_{cd}^{x_d} (1 - p_{cd})^{(1-x_d)}$
- C) $c^*(x) = \arg \max_{c=1, \dots, C} \log P(c) + \sum_{d=1}^D \log p_{cd}^{x_d} + \log(1 - p_{cd})^{x_d}$
- D) $c^*(x) = \arg \max_{c=1, \dots, C} \sum_{d=1}^D p_{cd}^{x_d} (1 - p_{cd})^{(1-x_d)}$

A) Sea una distribución multinomial de parámetro $\mathbf{p} = (\frac{1}{4}, \frac{1}{2}, 0, \frac{1}{8}, \frac{1}{8})$. Al aplicar descuento absoluto de valor $\epsilon = \frac{1}{16}$ y distribuir por *back-off*, ¿cuál es el prototipo multinomial resultante?

- A) $(\frac{3}{16}, \frac{7}{16}, \frac{4}{16}, \frac{1}{16}, \frac{1}{16})$
- B) $(\frac{19}{80}, \frac{39}{80}, \frac{4}{80}, \frac{9}{80}, \frac{9}{80})$
- C) $(\frac{5}{21}, \frac{9}{21}, \frac{1}{21}, \frac{3}{21}, \frac{3}{21})$
- D) $(\frac{4}{16}, \frac{8}{16}, \frac{1}{16}, \frac{2}{16}, \frac{2}{16})$

C) Con un clasificador gaussiano general se obtienen fronteras de decisión de tipo:

- A) Lineal
- B) Lineal a trozos
- C) Cuadráticas
- D) De otro tipo

D) A diferencia del clasificador de Bayes, un clasificador k -NN:

- A) Nunca puede ser óptimo
- B) No puede verse con un equivalente probabilístico
- C) Se puede aplicar también a datos no vectoriales
- D) Realiza estimaciones directas de $P(c|x)$

A) Si se aplica condensado sobre un conjunto no editado:

- A) Se pueden mantener prototipos fuera de norma (*outliers*)
- B) No se reducirá en ningún caso el conjunto de prototipos
- C) Se mantiene el coste computacional con respecto al que tendría con un conjunto editado
- D) No se mantienen los prototipos cercanos a las fronteras de decisión

Examen de Teoría de Percepción - Recuperación Segundo Parcial

ETSINF, Universitat Politècnica de València, Junio de 2019

Apellidos:

Nombre:

Profesor: ☐ Jorge Civera ☐ Carlos Martínez

Problemas (4 puntos, 90 minutos, con apuntes)

1. (0.5 puntos) Sea la función $K(\mathbf{x}, \mathbf{y}) = \frac{\mathbf{x}^t \cdot \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$, también conocida como distancia coseno entre dos vectores, y el conjunto de muestras

$$X = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$$

- a) Calcula la matriz Gramm \mathbf{K} asociada a las muestras de X con la función $K(\mathbf{x}, \mathbf{y})$ (0.3 puntos)
b) Indica si la función $K(\mathbf{x}, \mathbf{y})$ es un Kernel, justificando la respuesta (0.2 puntos)

Solución:

a)

$$\mathbf{K} = \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & -1 \\ \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \\ -1 & -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 1 \end{pmatrix}$$

- b) Sí es un Kernel, pues es un Kernel polinomial ($c = 0, d = 1$) multiplicado, en este caso, por una constante $c > 0$ que el producto de los módulos.

2. (1.5 puntos) Se tiene el siguiente conjunto de muestras bidimensionales de dos clases A y B :

n	1	2	3	4	5	6	7	8
x_1	3	3	5	5	3	0	0	3
x_2	2	4	2	4	0	2	4	6
c	A	A	A	A	B	B	B	B

Se pide:

- a) Estima los parámetros multinomiales para las clases A y B tomando como entrenamiento el conjunto de muestras presentado (0.3 puntos)
b) Estima los parámetros de las distribuciones gaussianas para las clases A y B (media μ_c y matriz de covarianzas Σ_c) tomando como entrenamiento el conjunto de muestras presentado, suponiendo que las matrices de covarianzas son distintas para cada clase (0.6 puntos)
c) Clasifica el dato $\mathbf{x} = (4, 3)^t$ con los clasificadores multinomial y gaussiano asociados a los parámetros inferidos en los apartados anteriores (0.6 puntos)

Nota: recuerda que la matriz de covarianzas se calcula como $\Sigma_c = \frac{1}{N_c} \sum_{n=1}^{N_c} (\mathbf{x}_n - \mu_c)(\mathbf{x}_n - \mu_c)^t$

Solución:

a) $\mathbf{p}_A = \left(\frac{4}{7}, \frac{3}{7}\right)$, $\mathbf{p}_B = \left(\frac{1}{3}, \frac{2}{3}\right)$

b) $\mu_A = (4, 3)^t$, $\mu_B = \left(\frac{3}{2}, 3\right)^t$, $\Sigma_A = I$, $\Sigma_B = \begin{pmatrix} \frac{9}{4} & 0 \\ 0 & 5 \end{pmatrix}$

- c) Como $P(A) = P(B)$ se puede ignorar en todos los clasificadores. Por tanto, para el clasificador multinomial:

$$g_A(\mathbf{x}) = \left(\frac{4}{7}\right)^4 \left(\frac{3}{7}\right)^3 = \frac{6912}{823543} \approx 0.00839 \quad g_B(\mathbf{x}) = \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^3 = \frac{8}{2187} \approx 0.00366$$

Con lo que se clasifica en la clase A.

Respecto al clasificador gaussiano, teniendo en cuenta que $\Sigma_A^{-1} = I$ y que $\Sigma_B^{-1} = \begin{pmatrix} \frac{4}{9} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}$:

$$g_A(\mathbf{x}) = |\Sigma_A|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (\mathbf{x} - \mu_A)^t \Sigma_A^{-1} (\mathbf{x} - \mu_A) \right) = \exp \left(-\frac{1}{2} \begin{pmatrix} 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \exp(0) = 1$$

$$g_B(\mathbf{x}) = |\Sigma_B|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} (\mathbf{x} - \mu_B)^t \Sigma_B^{-1} (\mathbf{x} - \mu_B) \right) = \frac{2}{3\sqrt{5}} \exp \left(-\frac{1}{2} \begin{pmatrix} \frac{5}{2} & 0 \end{pmatrix} \begin{pmatrix} \frac{4}{9} & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} \frac{5}{2} \\ 0 \end{pmatrix} \right) =$$

$$\frac{2}{3\sqrt{5}} \exp \left(-\frac{1}{2} \begin{pmatrix} \frac{10}{9} & 0 \end{pmatrix} \begin{pmatrix} \frac{5}{2} \\ 0 \end{pmatrix} \right) = \frac{2}{3\sqrt{5}} \exp \left(-\frac{25}{18} \right) \approx 0.074$$

Con lo que se clasifica en la clase A.

3. **(1 punto)** Tenemos el siguiente conjunto de prototipos bidimensionales de dos clases A y B:

\mathbf{x}_1	(-1,2)	A	\mathbf{x}_2	(-1,-2)	A
\mathbf{x}_3	(1,1)	B	\mathbf{x}_4	(1,-1)	B
\mathbf{x}_5	(3,1)	B	\mathbf{x}_6	(3,-1)	B
\mathbf{x}_7	(-2,0)	A	\mathbf{x}_8	(1,0)	A

Se pide:

- Clasificar la muestra $\mathbf{y} = (0,0)$ empleando distancia euclídea al cuadrado mediante vecino más cercano (NN) y 3 vecinos más cercanos (3-NN) **(0.3 puntos)**
- Aplicar una iteración del algoritmo de edición de Wilson en orden ascendente con la misma métrica y NN **(0.3 puntos)**
- Aplicar una iteración del algoritmo de edición de Wilson en orden ascendente con la misma métrica y 3-NN **(0.4 puntos)**

Nota: en caso de empates de distancias, suponed que la ganadora es la clase correcta.

Solución:

- Las distancias entre \mathbf{y} y los diversos prototipos son:

\mathbf{x}_1	\mathbf{x}_2	\mathbf{x}_3	\mathbf{x}_4	\mathbf{x}_5	\mathbf{x}_6	\mathbf{x}_7	\mathbf{x}_8
5	5	2	2	10	10	4	1

Por tanto, por NN se clasifica en la clase A (\mathbf{x}_8) mientras que por 3-NN se clasifica en la clase B ($\mathbf{x}_8, \mathbf{x}_3, \mathbf{x}_4$)

- La matriz de distancias entre los prototipos es la siguiente:

\mathbf{x}_1	\mathbf{x}_1							
\mathbf{x}_2	16	\mathbf{x}_2						
\mathbf{x}_3	5	13	\mathbf{x}_3					
\mathbf{x}_4	13	5	4	\mathbf{x}_4				
\mathbf{x}_5	17	25	4	8	\mathbf{x}_5			
\mathbf{x}_6	25	17	8	4	4	\mathbf{x}_6		
\mathbf{x}_7	5	5	10	10	26	26	\mathbf{x}_7	
\mathbf{x}_8	8	8	1	1	5	5	9	\mathbf{x}_8

Iterando en orden ascendente por NN:

- $\mathbf{x}_1 \rightarrow \mathbf{x}_3, \mathbf{x}_7 \rightarrow$ Empate, a correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_2 \rightarrow \mathbf{x}_4, \mathbf{x}_7 \rightarrow$ Empate, a correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_3 \rightarrow \mathbf{x}_8 \rightarrow$ incorrecta (A), se elimina ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_4 \rightarrow \mathbf{x}_8 \rightarrow$ incorrecta (A), se elimina ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_5 \rightarrow \mathbf{x}_6 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_6 \rightarrow \mathbf{x}_5 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_7 \rightarrow \mathbf{x}_1, \mathbf{x}_2 \rightarrow$ correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_8 \rightarrow \mathbf{x}_5, \mathbf{x}_6 \rightarrow$ incorrecta (B), se elimina ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$)

Por tanto, el conjunto final tras la primera iteración sería $X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$

- Iterando en orden ascendente por 3-NN:

- $\mathbf{x}_1 \rightarrow \mathbf{x}_3, \mathbf{x}_7, \mathbf{x}_8 \rightarrow$ correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_2 \rightarrow \mathbf{x}_4, \mathbf{x}_7, \mathbf{x}_8 \rightarrow$ correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_3 \rightarrow \mathbf{x}_8, \mathbf{x}_4, \mathbf{x}_5 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)

- $\mathbf{x}_4 \rightarrow \mathbf{x}_8, \mathbf{x}_4, \mathbf{x}_6 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_5 \rightarrow \mathbf{x}_3, \mathbf{x}_6, \mathbf{x}_8 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_6 \rightarrow \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_8 \rightarrow$ correcta (B), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_7 \rightarrow \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_8 \rightarrow$ correcta (A), permanece ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7, \mathbf{x}_8\}$)
- $\mathbf{x}_8 \rightarrow \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5/\mathbf{x}_6 \rightarrow$ incorrecta (B), se elimina ($X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$)

Por tanto, el conjunto final tras la primera iteración sería $X = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_7\}$

4. (1 punto) Sean las siguientes muestras y clasificadores:

$$\mathbf{x}_1 = (-2, -1) \in +1 \quad \mathbf{x}_2 = (-1, -2) \in -1 \quad \mathbf{x}_3 = (1, -2) \in +1 \quad \mathbf{x}_4 = (2, -1) \in -1$$

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

Aplica una iteración de AdaBoost para ese conjunto de datos y clasificadores indicando:

- Clasificador escogido C_1 .
- Valor de ϵ_1 .
- Valor de α_1 .
- Actualización de los pesos para la siguiente iteración ($w^{(2)}$).

Solución:

Tabla de acierto/fallo:

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

Pesos iniciales: $w^{(1)} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$

$$C_1 = g_3$$

Sumatorio de los $w^{(1)}$ de las muestras incorrectas:

$$\epsilon_1 = \frac{1}{4}$$

g_1	g_2	g_3	g_4
$\frac{2}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{3}{4}$

$$\alpha_1 = \frac{1}{2} \ln 3$$

	$w^{(1)} \exp(-y_i \alpha_1 C_1(x_i))$
\mathbf{x}_1	$\frac{1}{4} e^{-\frac{1}{2} \ln 3}$
\mathbf{x}_2	$\frac{1}{4} e^{-\frac{1}{2} \ln 3}$
\mathbf{x}_3	$\frac{1}{4} e^{-\frac{1}{2} \ln 3}$
\mathbf{x}_4	$\frac{1}{4} e^{-\frac{1}{2} \ln 3}$
Suma total	$\frac{\sqrt{3}}{2}$

$$w^{(2)} = (\frac{1}{6}, \frac{1}{6}, \frac{1}{2}, \frac{1}{6})$$