Examen de Teoría de Percepción - Recuperación Segundo Parcial

ETSINF, Universitat Politècnica de València, Junio de 2014

Apellidos: ______________________ **Nombre:** ______________

Profesor: ☐ **Jorge Civera** ☐ **Roberto Paredes**

Cuestiones (3 puntos, 30 minutos, sin apuntes)

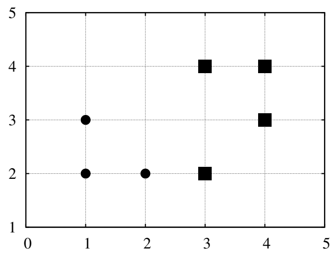
$\boxed{\text{B}}$ Sea una representación vectorial de los objetos en d dimensiones y un clasificador basado en distancias ¿Cuál de las siguientes afirmaciones sobre las distancias es cierta?

A) $L_2(\mathbf{x}, \mathbf{y}) = L_1(\mathbf{x}, \mathbf{y}) \ \ \forall \mathbf{x}, \mathbf{y}$
B) $L_2(\mathbf{x}, \mathbf{y}) \leq L_1(\mathbf{x}, \mathbf{y}) \ \ \forall \mathbf{x}, \mathbf{y}$
C) $L_1(\mathbf{x}, \mathbf{y}) \leq L_2(\mathbf{x}, \mathbf{y}) \ \ \forall \mathbf{x}, \mathbf{y}$
D) $L_1(\mathbf{x}, \mathbf{y}) \leq L_\infty(\mathbf{x}, \mathbf{y}) \ \ \forall \mathbf{x}, \mathbf{y}$

$\boxed{\text{C}}$ Sea una representación vectorial de los objetos en d dimensiones y un clasificador basado en distancias. El resultado de la regla de clasificación:

A) es la misma si se emplea la distancia L_2 o L_1
B) es la misma si se emplea con $k = 3$ o $k = 5$
C) en una dimensión es la misma si se emplea la distancia L_1 o L_∞
D) es la misma si se emplea la distancia L_2 o L_∞

$\boxed{\text{B}}$ La figura de la derecha muestra prototipos bidimensionales de 2 clases, $X = \{x_1 = (1, 2, \bullet),\ x_2 = (3, 2, \blacksquare),\ x_3 = (2, 2, \bullet),\ x_4 = (4, 3, \blacksquare),\ x_5 = (4, 4, \blacksquare),\ x_6 = (3, 4, \blacksquare), x_7 = (1, 3, \bullet)\}$. ¿Cuál es el conjunto de prototipos X' resultante de aplicar el algoritmo de edición de Wilson con el vecino más cercano en distancia Euclídea?



A) $X' = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$
B) $X' = \{x_1, x_3, x_4, x_5, x_6, x_7\}$
C) $X' = \{x_1, x_3, x_5, x_6, x_7\}$
D) $X' = \{x_1, x_3, x_5, x_7\}$

$\boxed{\text{C}}$ Considérese el conjunto de prototipos del ejercicio anterior. Si ejecutáramos el algoritmo de condensado de prototipos tras aplicar el algoritmo de edición de Wilson. ¿Qué conjunto de prototipos S obtendríamos como resultado del algoritmo de condensado con el vecino más cercano en distancia Euclídea?

A) $S = \{x_1, x_5, x_6\}$
B) $S = \{x_1, x_6\}$
C) $S = \{x_1, x_4\}$
D) $S = \{x_1, x_4, x_6\}$

$\boxed{\text{D}}$ Sean A, B , C y D cuatro clases con la misma probabilidad a priori y f.d. condicional de clase de tipo Bernoulli $\mathbf{p}_A = \begin{pmatrix} 0.4 \\ 0.9 \end{pmatrix}$, $\mathbf{p}_B = \begin{pmatrix} 0.4 \\ 0.5 \end{pmatrix}$, $\mathbf{p}_C = \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}$ y $\mathbf{p}_D = \begin{pmatrix} 0.2 \\ 0.6 \end{pmatrix}$. ¿En qué clase sería clasificada una muestra $\mathbf{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ dados los parámetros anteriores?

A) Clase A
B) Clase B
C) Clase C
D) Clase D

A Dados los parámetros Bernoulli de las clases del ejercicio anterior, ¿qué clases verían modificados los valores de sus parámetros si aplicáramos un suavizado por *Muestra ficticia*?

A) A,B,C,D
B) A,B
C) A,D
D) B,C

C Dado el siguiente conjunto de vectores enteros bidimensionales:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	3	2	1	2	3	1	1	1	0	0	0	0
x_{n2}	2	1	0	2	1	0	1	2	1	1	2	1
c_n	1	1	1	1	1	1	2	2	2	2	2	2

¿Cuál es la estimación de los parámetros del clasificador multinomial más probable?

A) $\hat{p}(1) = \hat{p}(2) = \frac{1}{4}$, $\hat{\mathbf{p}}_1 = \left(\frac{1}{3}, \frac{2}{3}\right)^t$ y $\hat{\mathbf{p}}_2 = \left(\frac{1}{5}, \frac{4}{5}\right)^t$
B) $\hat{p}(1) = \hat{p}(2) = \frac{1}{4}$, $\hat{\mathbf{p}}_1 = \left(\frac{2}{3}, \frac{1}{3}\right)^t$ y $\hat{\mathbf{p}}_2 = \left(\frac{4}{5}, \frac{1}{5}\right)^t$
C) $\hat{p}(1) = \hat{p}(2) = \frac{1}{2}$, $\hat{\mathbf{p}}_1 = \left(\frac{2}{3}, \frac{1}{3}\right)^t$ y $\hat{\mathbf{p}}_2 = \left(\frac{1}{5}, \frac{4}{5}\right)^t$
D) $\hat{p}(1) = \hat{p}(2) = \frac{1}{2}$, $\hat{\mathbf{p}}_1 = \left(\frac{1}{3}, \frac{2}{3}\right)^t$ y $\hat{\mathbf{p}}_2 = \left(\frac{2}{5}, \frac{3}{5}\right)^t$

C Sea un problema de clasificación en 10 clases donde los objetos se representan en 100 dimensiones. Las f.d. condicional de cada clase se modelan mediante mixturas de 8 gausianas con matrices de covarianzas diagonales. Cuantos bytes son necesarios para almacenar todos los parámetros del clasificador si empleamos *float* (4 bytes) como precisión numérica.

A) Menos de 50 Kbytes
B) Entre 50Kbytes y 60Kbytes
C) Entre 60Kbytes y 70Kbytes
D) Más de 70 Kbytes

C Cuál de las siguientes afirmaciones sobre el aprendizaje de los parámetros de las mixturas de gaussianas es correcta:

A) Se realiza por conteo en una única iteración
B) Se realiza por conteo simple como las Bernoullis
C) Se realiza mediante el algoritmo EM
D) Se realiza por conteo simple como las Multinomiales

A En interacción hombre-máquina, cuál de los siguientes modelos de interacción resulta en general más sencillo de modelar:

A) Realimentación unimodal determinista
B) Realimentación multi-modal no-determinista
C) Realimentación unimodal no-determinista
D) Realimentación multi-modal determinista

Examen de Teoría de Percepción - Recuperación Segundo Parcial

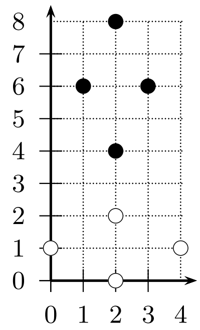
ETSINF, Universitat Politècnica de València, Junio de 2014

Apellidos: ☐ **Nombre:** ☐

Profesor: ☐ **Jorge Civera** ☐ **Roberto Paredes**

Problemas (4 puntos, 90 minutos, con apuntes)

1. **(2 puntos)** La figura de la derecha muestra prototipos bidimensionales de 2 clases: $\{x_1 = (2, 0, \circ), x_2 = (0, 1, \circ), x_3 = (4, 1, \circ), x_4 = (2, 2, \circ), x_5 = (2, 4, \bullet), x_6 = (1, 6, \bullet), x_7 = (3, 6, \bullet), x_8 = (2, 8, \bullet)\}$ utilizados en un clasificador por el vecino más cercano ($k = 1$). Se pide (0.5 puntos por apartado):



   a) Clasificar las muestras $\{y_1 = (2, 3),\ y_2 = (1, 3),\ y_3 = (0, 3)\}$ por distancia Euclídea.

   b) Clasificar las muestras del apartado a) por distancia Mahalanobis-diagonal por clase.

   c) Identificar dos puntos de la frontera de decisión asociada al clasificador de distancia Euclídea.

   d) Ídem al apartado c) pero para el clasificador de distancia Mahalanobis-diagonal por clase.

Solución:

a) La clasificación basada en el vecino más cercano por distancia Euclídea

$$d_E(\mathbf{y}, \mathbf{p}) = \sqrt{\sum_{i=1}^{D} (y_i - p_i)^2}$$

resulta en un empate por distancia para las muestras y_1 y y_2, así que se clasifican en cualquiera de las dos clases, mientras que $\hat{c}(y_3) = \circ$.

b) La clasificación basada en el vecino más cercano por distancia Mahalanobis-diagonal por clase

$$d_M(\mathbf{y}, \mathbf{p}) = \sqrt{\sum_{i=1}^{D} \frac{1}{\sigma_{ic}^2} (y_i - p_i)^2}$$

donde c es la clase de $\mathbf{p}$, y σ_{ic}^2 es la varianza de la componente i-ésima en la clase c. Para calcular la distancia de Mahalanobis-diagonal por clase entre cada prototipo $\mathbf{p}$ y una muestra $\mathbf{y}$, necesitamos calcular previamente las varianzas por clase y componente:

σ_{ic}^2	$i = 1$	$i = 2$
$c = \circ$	2	0.5
$c = \bullet$	0.5	2

Teóricamente para calcular el vecino más cercano deberíamos calcular la distancia de cada prototipo a la muestra $\mathbf{y}$ a clasificar. En esta solución, calcularemos la distancia para el prototipo de cada clase que es el más cercano a la muestra $\mathbf{y}$, pero en la práctica se calcularía para cualquier prototipo que tuviera opciones de ser el más cercano a la muestra $\mathbf{y}$. Para la muestra y_1:

$$d_M(y_1, x_4) = \sqrt{\frac{1}{2}(2-2)^2 + \frac{1}{0.5}(3-2)^2} = \sqrt{2}$$

$$d_M(y_1, x_5) = \sqrt{\frac{1}{0.5}(2-2)^2 + \frac{1}{2}(3-4)^2} = \sqrt{0.5}$$

Así que $\hat{c}(y_1) = \bullet$. Para la muestra y_2:

$$d_M(y_2, x_4) = \sqrt{\frac{1}{2}(1-2)^2 + \frac{1}{0.5}(3-2)^2} = \sqrt{2.5}$$

$$d_M(y_2, x_5) = \sqrt{\frac{1}{0.5}(1-2)^2 + \frac{1}{2}(3-4)^2} = \sqrt{2.5}$$

Se produce un empate por distancia clasificándose en cualquiera de las dos clases. Para la muestra y_3:

$$d_M(y_3, x_4) = \sqrt{\frac{1}{2}(0-2)^2 + \frac{1}{0.5}(3-2)^2} = \sqrt{4}$$

$$d_M(y_3, x_5) = \sqrt{\frac{1}{0.5}(0-2)^2 + \frac{1}{2}(3-4)^2} = \sqrt{8.5}$$

Así que $\hat{c}(y_3) = \circ$.

c) Los puntos de la frontera de decisión entre las dos clases son los puntos del espacio donde se producen empates. Así que aprovechando el resultado del apartado a) estos puntos serían $(2, 3)$ y $(1, 3)$.

d) Al igual que en el apartado anterior sabemos que $(1, 3)$ es un punto de la frontera de decisión. Dada la simetría en la distribución de los prototipos, podemos asegurar que otro punto en la frontera de decisión será $(3, 3)$.

2. **(2 puntos)** Tenemos $N = 12$ vectores 4-dimensionales etiquetados:

n	1	2	3	4	5	6	7	8	9	10	11	12
x_{n1}	1	0	1	1	0	1	0	0	0	0	0	0
x_{n2}	0	1	1	1	0	1	0	0	0	0	0	0
x_{n3}	0	0	0	0	0	0	1	1	0	1	0	1
x_{n4}	0	0	0	0	0	0	1	1	1	1	1	1
c_n	1	1	1	1	1	1	2	2	2	2	2	2

a) Calcula los parámetros del clasificador Bernoulli más probable respecto a estos datos, asumiendo que han sido aleatoriamente extraídos de $C = 2$ distribuciones Bernoulli independientes. (0.5 puntos)

b) Suaviza los parámetros Bernoulli del apartado anterior mediante truncamiento simple de $\epsilon = 1/3$. (0.5 puntos)

c) Calcula los parámetros del clasificador multinomial más probable respecto a estos datos, asumiendo que han sido aleatoriamente extraídos de $C = 2$ distribuciones multinomiales independientes. (0.5 puntos)

d) Suaviza los parámetros multinomiales del apartado anterior con descuento absoluto de $\epsilon = 0.2$ e interpolación usando como distribución generalizada, la distribución uniforme. (0.5 puntos)

Solución:

a)

$$p(1) = p(2) = \frac{6}{12} = 0.5$$

$$\hat{\mathbf{p}}_1 = \frac{1}{6}\begin{pmatrix} 1+0+1+1+0+1 \\ 0+1+1+1+0+1 \\ 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 4 \\ 4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 2/3 \\ 0.0 \\ 0.0 \end{pmatrix}$$

$$\hat{\mathbf{p}}_2 = \frac{1}{6}\begin{pmatrix} 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \\ 1+1+0+1+0+1 \\ 1+1+1+1+1+1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 0 \\ 0 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 2/3 \\ 1.0 \end{pmatrix}$$

b)

$$\hat{\mathbf{p}}_1 = \begin{pmatrix} 2/3 \\ 2/3 \\ 0.0 \\ 0.0 \end{pmatrix} \rightarrow \hat{\mathbf{p}}_1 = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

$$\hat{\mathbf{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 2/3 \\ 1.0 \end{pmatrix} \rightarrow \hat{\mathbf{p}}_2 = \begin{pmatrix} 1/3 \\ 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}$$

c)

$$p(1) = p(2) = \frac{6}{12} = 0.5$$

$$\hat{\mathbf{p}}_1 = \frac{1}{8}\begin{pmatrix} 1+0+1+1+0+1 \\ 0+1+1+1+0+1 \\ 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \end{pmatrix} = \frac{1}{8}\begin{pmatrix} 4 \\ 4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.0 \\ 0.0 \end{pmatrix}$$

$$\hat{\mathbf{p}}_2 = \frac{1}{10}\begin{pmatrix} 0+0+0+0+0+0 \\ 0+0+0+0+0+0 \\ 1+1+0+1+0+1 \\ 1+1+1+1+1+1 \end{pmatrix} = \frac{1}{10}\begin{pmatrix} 0 \\ 0 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.4 \\ 0.6 \end{pmatrix}$$

d)

$$\hat{\mathbf{p}}_1 = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.0 \\ 0.0 \end{pmatrix} \rightarrow \hat{\mathbf{p}}_1 = \begin{pmatrix} 0.5 - 0.2 + \frac{1}{4} \cdot 0.4 \\ 0.5 - 0.2 + \frac{1}{4} \cdot 0.4 \\ 0.0 \qquad + \frac{1}{4} \cdot 0.4 \\ 0.0 \qquad + \frac{1}{4} \cdot 0.4 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.4 \\ 0.1 \\ 0.1 \end{pmatrix}$$

$$\hat{\mathbf{p}}_2 = \begin{pmatrix} 0.0 \\ 0.0 \\ 0.4 \\ 0.6 \end{pmatrix} \rightarrow \hat{\mathbf{p}}_2 = \begin{pmatrix} 0.0 \qquad + \frac{1}{4} \cdot 0.4 \\ 0.0 \qquad + \frac{1}{4} \cdot 0.4 \\ 0.4 - 0.2 + \frac{1}{4} \cdot 0.4 \\ 0.6 - 0.2 + \frac{1}{4} \cdot 0.4 \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.1 \\ 0.3 \\ 0.5 \end{pmatrix}$$