Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Marzo de 2018

Apellidos: ______ **Nombre:** ______

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

Cuestiones (2 puntos, 30 minutos, sin apuntes)

$\boxed{\text{A}}$ Indica cuál de los siguientes es un clasificador de mínimo error para un objeto x sobre un conjunto de clases $\mathbb{C}$:

A) $\hat{c}(x) = \arg\max_{c\in\mathbb{C}} p(x, c)$
B) $\hat{c}(x) = \arg\min_{c\in\mathbb{C}} p(x, c)$
C) $\hat{c}(x) = \arg\max_{c\in\mathbb{C}} p(x, c) \cdot P(c)$
D) $\hat{c}(x) = \arg\max_{c\in\mathbb{C}} \frac{p(x,c)}{P(c)}$

$\boxed{\text{C}}$ En el esquema de clasificación con realimentación y reentrenamiento:

A) Nunca hay modelo inicial de clasificación.
B) Se emplea sólo la muestra y el resultado del clasificador para cambiar el modelo de clasificación.
C) Se debe decidir la estrategia de modificación del modelo según la realimentación recibida.
D) Se emplea sólo la realimentación humana para cambiar el modelo de clasificación.

$\boxed{\text{B}}$ Ante una representación por características locales que emplea ventanas de 5×5 píxeles sobre imágenes de 64 niveles de gris, ¿qué afirmación es correcta?

A) El tamaño en memoria de cada ventana por representación por histograma será de 25 bytes
B) La memoria total ocupada no puede calcularse sólo con estos datos
C) La memoria total ocupada usando representación directa por cada ventana será superior a 1000 bytes
D) El tamaño en memoria de cada ventana por representación directa será de 12 bytes

$\boxed{\text{D}}$ ¿Cuántos *frames* F se extraen de un segundo de señal de 16kHz con un desplazamiento de 100 muestras?

A) $0 < F \leq 50$ frames
B) $50 < F \leq 100$ frames
C) $100 < F \leq 150$ frames
D) $150 < F \leq 200$ frames

$\boxed{\text{B}}$ En la representación secuencia de tokens (n-gramas), al aumentar el valor de n su requerimiento de memoria crece:

A) De forma lineal con el tamaño del vocabulario.
B) De forma exponencial con el tamaño del vocabulario.
C) De forma logarítmica con el tamaño del vocabulario.
D) De forma constante porque no depende de la talla del vocabulario.

$\boxed{\text{C}}$ Sea un problema de clasificación en dos clases sobre vectores de $\mathbb{R}^4$, con las muestras $\{(0,0,1,1),\ (1,0,1,1),\ (0,1,1,0)\}$ para la clase 1 y las muestras {(1,1,1,0),(2,0,0,1),(1,0,1,2)} para la clase 2. Indicar qué matriz de proyección para $\mathbb{R}^2$ es la más apropiada para estas muestras.

A) $\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$
B) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$
C) $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
D) $\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

$\boxed{\text{A}}$ Indicar cuál de las siguientes es una característica de PCA.

A) Se basa en minimizar el error de reconstrucción.
B) Es una técnica supervisada.
C) La proyección resultado garantiza el mínimo error de clasificación.
D) Construye la matriz de proyección con los vectores propios de la matriz entre clases.

$\boxed{\text{D}}$ Dado el siguiente conjunto de muestras etiquetadas $\{(-2\quad 0),\ (-1\quad 0)\}$ de la clase A y $\{(1\quad 0),\ (2\quad 0)\}$ de la clase B en $\mathbb{R}^2$, ¿qué vector de proyección a una única dimensión no esperarías que resultara de aplicar LDA a estas muestras?

A) $\mathbf{w} = (1\quad 0)$
B) $\mathbf{w} = (1\quad 1)$
C) $\mathbf{w} = (-1\quad 1)$
D) $\mathbf{w} = (0\quad 1)$

Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Marzo de 2018

Apellidos: ____________________ **Nombre:** __________

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

Problemas (4 puntos, 90 minutos, con apuntes)

1. Calcula espacio requerido en memoria de las siguientes representaciones:

   a) Representación de una imagen de 27×50 píxeles en escala de grises de 256 niveles mediante características locales de 7×7 píxeles extraídas una cada 3 píxeles en horizontal y una cada 4 píxeles en vertical (**0.25 puntos**).
   Solución: $\frac{(27-7+1)}{3} \cdot \frac{(50-7+1)}{4} \cdot 7 \cdot 7 = 3773$ bytes

   b) Representación de una imagen de 60×60 píxeles en escala de grises de 256 niveles mediante características locales de 11×11 píxeles extraídas una cada 5 píxeles en ambas dimensiones y representada cada una de ellas mediante histograma (**0.25 puntos**).
   Solución: $\frac{(60-11+1)}{5} \cdot \frac{(60-11+1)}{5} \cdot 256 = 25600$ bytes

   c) Representación de la señal de audio en estéreo de un vídeo de Polimedia grabado a 16KHz que tiene un duración de 10 minutos donde cada muestra se representa mediante 16 bits (**0.25 puntos**).
   Solución: $10 \cdot 60 \cdot 16000 \cdot 2 \cdot 2 = 36.62$ Mbytes

   d) Representación de la señal de audio del apartado anterior habiendo sido procesada para mono (1 canal) y almacenada mediante sus frames con un tamaño $W = 25$ ms y desplazamiento $S = 10$ ms. Considera que cada componente del frame se almacena mediante un *float* de 4 bytes (**0.25 puntos**).
   Solución: $\frac{10 \cdot 60 \cdot 1000}{10} \cdot \left(25 \cdot \frac{16000}{1000}\right) \cdot 4 = 91.55$ Mbytes

2. Asume que se dispone de una colección de $D > 0$ documentos con un token t_1 que ocurre con una frecuencia constante $k > 0$ en todos los documentos, y un token t_2 que ocurre únicamente en uno de los documentos con frecuencia $k > 0$. Se pide:

   a) Calcular las funciones globales *Normal*, *GfIdf* y *Idf* para el token t_1 y discutir qué función global le asignaría un menor valor (**0.5 puntos**).

   b) Calcular las funciones globales *Normal*, *GfIdf* y *Idf* para el token t_2 y discutir qué función global le asignaría un menor valor (**0.5 puntos**).

   Solución:

   a) Sustituimos en las funciones globales asumiendo la frecuencia de t_1:

$$\text{Normal} \qquad G(t_1) = \left(\sum_d x_{dt_1}^2\right)^{-\frac{1}{2}} = \left(\sum_d k^2\right)^{-\frac{1}{2}} = \left(D \cdot k^2\right)^{-\frac{1}{2}}$$

$$\text{GfIdf} \qquad G(t_1) = \frac{\sum_d x_{dt_1}}{\sum_{d:x_{dt_1}>0} 1} = \frac{\sum_d k}{D} = \frac{D \cdot k}{D} = k$$

$$\text{Idf} \qquad G(t_1) = \log \frac{D}{\sum_{d:x_{dt_1}>0} 1} = \log \frac{D}{D} = 0$$

   siendo la función global *Idf* la de menor valor.

   b) Sustituimos en las funciones globales asumiendo la frecuencia de t_2:

$$\text{Normal} \qquad G(t_2) = \left(\sum_d x_{dt_2}^2\right)^{-\frac{1}{2}} = \left(k^2\right)^{-\frac{1}{2}} = \frac{1}{k}$$

$$\text{GfIdf} \qquad G(t_2) = \frac{\sum_d x_{dt_2}}{\sum_{d:x_{dt_2}>0} 1} = \frac{k}{1} = k$$

$$\text{Idf} \qquad G(t_2) = \log \frac{D}{\sum_{d:x_{dt_2}>0} 1} = \log \frac{D}{1} = \log D$$

En este caso, la función global de menor peso dependerá del valor de k y D. Asumiendo una colección de cientos de documentos, la función global *Normal* será menor o igual ($k = 1$) que las otras dos funciones.

3. Se tiene el siguiente conjunto de datos para dos clases en $\mathbb{R}^5$:

Muestra						Clase
x_1	4	1	-1	1	0	A
x_2	2	-1	1	0	1	A
x_3	4	5	-2	3	-1	A
x_4	-2	2	-3	-1	-2	B
x_5	2	-7	5	2	-3	B

Se ha empleado PCA para hallar los valores y vectores propios correspondientes, que se muestran en la siguiente tabla redondeados a dos decimales:

w_1	0.00	0.82	-0.56	-0.03	0.12	23.64
w_2	0.83	0.15	0.23	0.45	0.18	6.85
w_3	0.13	-0.14	0.00	-0.52	0.84	2.06
w_4	-0.48	0.41	0.64	0.30	0.32	0.25
w_5	-0.25	-0.35	-0.47	0.66	0.39	0.00

Se pide:

a) Calcular la proyección PCA a $\mathbb{R}^2$ de las muestras dadas (**0.75 puntos**).

b) Calcular las matrices S_b y S_w que se emplearían para calcular la proyección LDA a partir de las muestras proyectadas (de $\mathbb{R}^2$) del apartado previo (**1 punto**).

c) Si se quiere emplear un clasificador lineal, ¿bastaría con la proyección PCA propuesta o sería necesario aplicar LDA? Razona la respuesta (**0.25 puntos**).

Solución:

a) En primer lugar se calcula la media ($\bar{\mathbf{x}} = (2, 0, 0, 1, -1)$) y se resta a cada una de las muestras:

Muestra						Clase
$x_1 - \bar{\mathbf{x}}$	2	1	-1	0	1	A
$x_2 - \bar{\mathbf{x}}$	0	-1	1	-1	2	A
$x_3 - \bar{\mathbf{x}}$	2	5	-2	2	0	A
$x_4 - \bar{\mathbf{x}}$	-4	2	-3	-2	-1	B
$x_5 - \bar{\mathbf{x}}$	0	-7	5	1	-2	B

La matriz de proyección se define por los dos primeros vectores propios:

$$W^t = \begin{pmatrix} 0.00 & 0.82 & -0.56 & -0.03 & 0.12 \\ 0.83 & 0.15 & 0.23 & 0.45 & 0.18 \end{pmatrix}$$

Aplicando la proyección:

Muestra		
$W^t \cdot (x_1 - \bar{\mathbf{x}})$	1.50	1.76
$W^t \cdot (x_2 - \bar{\mathbf{x}})$	-1.11	-0.01
$W^t \cdot (x_3 - \bar{\mathbf{x}})$	5.16	2.85
$W^t \cdot (x_4 - \bar{\mathbf{x}})$	3.26	-4.79
$W^t \cdot (x_5 - \bar{\mathbf{x}})$	-8.81	0.19

b) Para calcular S_b es preciso calcular la media global de los datos proyectados $\bar{\mathbf{x}}^t = (0, 0)$ y la media por clase:

$$\bar{\mathbf{x}}_A = \begin{pmatrix} 1.85 \\ 1.53 \end{pmatrix} \qquad \bar{\mathbf{x}}_B = \begin{pmatrix} -2.78 \\ -2.30 \end{pmatrix}$$

A partir de aquí, como $S_b = \sum_{c=1}^{C} n_c(\bar{\mathbf{x}}_c - \bar{\mathbf{x}}) \cdot (\bar{\mathbf{x}}_c - \bar{\mathbf{x}})^t$, tendremos:

$$S_b = 3 \cdot \begin{pmatrix} 1.85 \\ 1.53 \end{pmatrix} \cdot (1.85, 1.53) + 2 \cdot \begin{pmatrix} -2.78 \\ -2.30 \end{pmatrix} \cdot (-2.78 - 2.30) = \begin{pmatrix} 25.613 & 21.233 \\ 21.233 & 17.603 \end{pmatrix}$$

Para calcular S_w es preciso calcular las matrices de covarianzas de los datos proyectados (x_i'), que son las siguientes:

$$\Sigma_A = \frac{1}{3}\left((x_1' - \bar{\mathbf{x}}_A) \cdot (x_1' - \bar{\mathbf{x}}_A)^t + (x_2' - \bar{\mathbf{x}}_A) \cdot (x_2' - \bar{\mathbf{x}}_A)^t + (x_3' - \bar{\mathbf{x}}_A) \cdot (x_3' - \bar{\mathbf{x}}_A)^t\right) =$$

$$\frac{1}{3}\left(\left(\begin{array}{c}-0.35\\0.23\end{array}\right)\cdot(-0.35,0.23)+\left(\begin{array}{c}-2.96\\-1.54\end{array}\right)\cdot(-2.96,-1.54)+\left(\begin{array}{c}3.31\\1.32\end{array}\right)\cdot(3.31,1.32)\right)=\left(\begin{array}{cc}6.61&2.95\\2.95&1.39\end{array}\right)$$

$$\Sigma_B=\frac{1}{2}\left((x'_4-\bar{\mathbf{x}}_B)\cdot(x'_4-\bar{\mathbf{x}}_B)^t+(x'_5-\bar{\mathbf{x}}_B)\cdot(x'_5-\bar{\mathbf{x}}_B)^t\right)=$$

$$\frac{1}{2}\left(\left(\begin{array}{c}6.04\\-2.49\end{array}\right)\cdot(6.04,-2.49)+\left(\begin{array}{c}-6.04\\2.49\end{array}\right)\cdot(-6.04,2.49)\right)=\left(\begin{array}{cc}36.42&-15.03\\-15.03&6.20\end{array}\right)$$

Por tanto:

$$S_w=\Sigma_A+\Sigma_B=\left(\begin{array}{cc}43.03&-12.08\\-12.08&7.59\end{array}\right)$$

c) No sería necesario proyectar a LDA; si tomamos la proyección PCA, se puede definir una frontera de decisión lineal entre ambas clases empleando, por ejemplo, $x_1+x_2=-1.5$ como frontera; de esta forma, aquellas muestras cuyas componentes sumen más de -1.5 quedan para la clase A y las que suman menos quedan para la clase B.