

Examen de Teoría de Percepción - Recuperación Primer Parcial

ETSINF, Universitat Politècnica de València, Junio de 2013

Apellidos:

Nombre:

Profesor: ☐ Carlos Martínez - Jorge Civera ☐ Roberto Paredes

Cuestiones (3 puntos, 30 minutos, sin apuntes)

☐ B ¿Cuál de los siguientes clasificadores **no** es un clasificador de mínimo riesgo o clasificador de Bayes?

- A) $c(x) = \arg \max_c \log(p(x, c))$
- B) $c(x) = \arg \max_c \log(p(x|c))$
- C) $c(x) = \arg \max_c \log(p(c|x))$
- D) $c(x) = \arg \min_c -\log(p(c|x))$

☐ C Dadas las funciones discriminantes $g_1(\mathbf{x}) = 3x_1 - 2x_2 + 1$ y $g_2(\mathbf{x}) = -x_1 + 2x_2 + 3$, ¿cuál es la frontera de decisión asociada?

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

☐ D Supongamos una representación local de imágenes de 256 niveles de gris de tamaño 16×16 píxeles, empleando ventanas de 5×5 ; en ese caso, la representación de la imagen ocupará:

- A) Menos de 1000 bytes.
- B) Entre 1000 y 1999 bytes.
- C) Entre 2000 y 2999 bytes.
- D) 3000 o más bytes.

☐ C El banco de filtros de Mel se aplica en reconocimiento de habla para:

- A) Pasar del dominio temporal al dominio frecuencial.
- B) Obtener los marcos (*frames*) de la señal acústica.
- C) Modelar la percepción de oído humano.
- D) Poder aplicar el reconocimiento del habla en tándem.

☐ A Indicar cual de los siguientes procesos **no** se aplica generalmente en reconocimiento del habla continua:

- A) Segmentación de traza.
- B) Preénfasis.
- C) Transformada de Fourier (FFT).
- D) Transformada del coseno (DCT).

☐ D ¿Qué objetivo persigue la función de ponderación global Idf ?

- A) Enfatizar aquellos tokens que ocurren en muchos documentos.
- B) Atenuar aquellos tokens con baja frecuencia.
- C) Enfatizar aquellos tokens con alta frecuencia.
- D) Atenuar aquellos tokens que ocurren en muchos documentos.

B Dadas las representaciones *bag-of-words* y bigramas de un conjunto de correos electrónicos. En general, se puede afirmar que:

- A) La matriz *bag-of-words* resultante será más dispersa que la matriz de bigramas.
- B) La matriz *bag-of-words* resultante será menos dispersa que la matriz de bigramas.
- C) La matriz *bag-of-words* resultante será tan dispersa como la matriz de bigramas.
- D) Ninguna de las anteriores.

A ¿Cuál de estas afirmaciones sobre PCA **no** es correcta?

- A) En el espacio proyectado mediante PCA se maximiza la correlación de las componentes.
- B) El objetivo de PCA es encontrar una proyección lineal que minimice el error de reconstrucción.
- C) La matriz de proyección lineal está compuesta por los eigenvectores con mayor eigenvalor asociado.
- D) Los eigenvectores son ortogonales entre sí.

B Dada la descomposición \mathbb{R}^3 en valores y vectores propios $\lambda_1 = 0.7$ con $\mathbf{w}_1 = (1 \ 0 \ 0)$, $\lambda_2 = 5.2$ con $\mathbf{w}_2 = (0 \ 1 \ 0)$, y $\lambda_3 = 2.7$ con $\mathbf{w}_3 = (0 \ 0 \ 1)$:

- A) La proyección PCA de \mathbb{R}^3 a \mathbb{R}^2 se llevará a cabo con los vectores propios \mathbf{w}_1 y \mathbf{w}_2 .
- B) La proyección PCA de \mathbb{R}^3 a \mathbb{R}^2 se llevará a cabo con los vectores propios \mathbf{w}_2 y \mathbf{w}_3 .
- C) La proyección PCA de \mathbb{R}^3 a \mathbb{R}^1 se llevará a cabo con el vector propio \mathbf{w}_1 .
- D) La proyección PCA de \mathbb{R}^3 a \mathbb{R}^1 se llevará a cabo con el vector propio \mathbf{w}_3 .

A En una proyección LDA donde el número de datos totales N es mayor que el número de clases C , ¿por qué está limitado a $C - 1$ el número máximo de dimensiones a las cuales se pueden proyectar los datos?

- A) Porque el rango de la matriz $S_b = \sum_{i=1}^C (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^t$ es como máximo $C - 1$.
- B) Porque el rango de la matriz $S_w = \sum_{i=1}^C \Sigma_i$ es $C - 1$.
- C) Porque el rango de la matriz S_w^{-1} es $C - 1$.
- D) Porque el número máximo de dimensiones a las cuales se pueden proyectar es siempre $C - 1$.

Examen de Teoría de Percepción - Recuperación Primer Parcial

ETSINF, Universitat Politècnica de València, Junio de 2013

Apellidos:

Nombre:

Profesor: ☐ Carlos Martínez - Jorge Civera ☐ Roberto Paredes

Problemas (4 puntos, 90 minutos, con apuntes)

1. (1 punto) Se tiene un clasificador en dos clases basado en distribuciones Bernoulli bidimensionales, de forma que para la clase 1 se tiene $p_1 = [0.3 \ 0.2]$, y para la clase 2 se tiene $p_2 = [0.6 \ 0.8]$. Se pide clasificar la muestra $y = [0 \ 1]$ empleando $\arg \max_c P(c | y)$ sabiendo que la probabilidad condicional $p(y | c) = \prod_d (p_{cd} y_d + (1 - p_{cd})(1 - y_d))$ y las probabilidades a priori son idénticas para ambas clases.

Solución:

$$\hat{c}(y) = \arg \max_c P(c | y) \approx \arg \max_c p(y | c) p(c)$$

Para la clase 1:

$$\begin{aligned} p(y = [0 \ 1] | c = 1) &= (p_{11} y_1 + (1 - p_{11})(1 - y_1)) \cdot (p_{12} y_2 + (1 - p_{12})(1 - y_2)) \\ &= (0.3 \cdot 0 + (1 - 0.3) \cdot (1 - 0)) \cdot (0.2 \cdot 1 + (1 - 0.2) \cdot (1 - 1)) \\ &= 0.14 \end{aligned}$$

Para la clase 2:

$$\begin{aligned} p(y = [0 \ 1] | c = 2) &= (p_{21} y_1 + (1 - p_{21})(1 - y_1)) \cdot (p_{22} y_2 + (1 - p_{22})(1 - y_2)) \\ &= (0.6 \cdot 0 + (1 - 0.6)(1 - 0)) \cdot (0.8 \cdot 1 + (1 - 0.8) \cdot (1 - 1)) \\ &= 0.32 \end{aligned}$$

Dado que $p(c = 1) = p(c = 2) = 0.5$, la muestra y se clasifica en la clase 2.

2. (1 punto) Calcular el tamaño en bytes que ocuparía cada una de estas señales acústicas, adquiridas en las condiciones indicadas:

- a) 3 minutos de una señal telefónica, adquirida a 8 KHz con 8 bits por muestra.
- b) 10 segundos de señal vocal captada por micrófono, adquirida a 16 KHz con 16 bits por muestra.
- c) 1 minuto de una señal de alta fidelidad para un sistema de audio 5.1, adquirida a 44 KHz con 16 bits por muestra.

Solución:

- a) $3 \cdot 60 \cdot 8000 \cdot 1 = 1440000$ bytes
- b) $10 \cdot 16000 \cdot 2 = 320000$ bytes
- c) $1 \cdot 60 \cdot 6 \cdot 44000 \cdot 2 = 31680000$ bytes

3. (2 puntos) Dada la siguiente matriz de proyección:

$$W = \begin{bmatrix} 1 & 1 & -1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

y las siguientes muestras etiquetadas $X = \{(\mathbf{x}_1, A), (\mathbf{x}_2, A), (\mathbf{x}_3, B), (\mathbf{x}_4, B)\}$ se pide obtener el error de clasificación en el espacio proyectado al emplear las siguientes funciones discriminantes lineales:

$$g_A(\mathbf{x}) = \mathbf{w}_A \mathbf{x}$$

$$g_B(\mathbf{x}) = \mathbf{w}_B \mathbf{x}$$

donde:

$$\mathbf{x}_1 = (1, 0, -1, 1), \mathbf{x}_2 = (1, 1, -2, 0), \mathbf{x}_3 = (-1, 2, 2, -1), \mathbf{x}_4 = (-1, 2, 2, -2) \text{ y}$$

$$\mathbf{w}_A = (1, 2, 1), \mathbf{w}_B = (1, -1, -1)$$

Solución: La proyección resultaría:

$$\mathbf{x}'_1 = W\mathbf{x}_1 = (3, 2)$$

$$\mathbf{x}'_2 = W\mathbf{x}_2 = (4, 1)$$

$$\mathbf{x}'_3 = W\mathbf{x}_3 = (-2, -2)$$

$$\mathbf{x}'_4 = W\mathbf{x}_4 = (-3, -3)$$

Para clasificar hay que emplear la notación compacta:

$$\mathbf{x}'_1 = (1, 3, 2)$$

$$\mathbf{x}'_2 = (1, 4, 1)$$

$$\mathbf{x}'_3 = (1, -2, -2)$$

$$\mathbf{x}'_4 = (1, -3, -3)$$

Y aplicar las funciones discriminantes de cada clase:

$$g_A(\mathbf{x}'_1) = 9 \quad g_B(\mathbf{x}'_1) = -4 \quad OK$$

$$g_A(\mathbf{x}'_2) = 10 \quad g_B(\mathbf{x}'_2) = -4 \quad OK$$

$$g_A(\mathbf{x}'_3) = -5 \quad g_B(\mathbf{x}'_3) = 5 \quad OK$$

$$g_A(\mathbf{x}'_4) = -8 \quad g_B(\mathbf{x}'_4) = 7 \quad OK$$

Por lo tanto se consigue un 0% de error.