Examen de Teoría de Percepción - Recuperación Primer Parcial

ETSINF, Universitat Politècnica de València, Junio de 2014

Apellidos: ____________ **Nombre:** ____________

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

Cuestiones (3 puntos, 30 minutos, sin apuntes)

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

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

B Dado un conjunto de muestras etiquetadas $X = \{(\mathbf{x}_1, c_1), \cdots, (\mathbf{x}_n, c_n)\}$ donde $\mathbf{x}_i \in \mathbb{R}^2$, podremos decir que la frontera de decisión entre dos clases es:

A) Un punto
B) Una línea
C) Una superfície
D) Ninguna de las anteriores

C Dada una representación local con un tamaño de ventana de 7x7 de un conjunto de 100 imágenes de 21x21 píxeles en 1024 niveles de gris, ¿cuál sería el espacio máximo que ocuparía este conjunto con esta representación local?

A) No más de 1 Mbyte
B) No menos de 5 Mbytes
C) Entre 2 y 3 MBytes
D) Ninguna de las anteriores

n = (21 - 7 + 1) * (21 - 7 + 1) = 225 -> 100 imágenes -> 225 * 100 = 22500

Representación directa -> n * (log2(niveles de gris) / 8)) -> log2(niveles de gris) / 8 = **log2(1024) / 8 = 1,25 -> se redondea a 2 -> Representación directa -> nd * 2 = 7 * 7 * 2 = 98 = s.**

Tamaño final = n * s = 22500 * 98 / 1000000 = 2.2MB.

A ¿Cuánto espacio como mínimo requiere el almacenamiento de una señal de 10 segundos con una frecuencia máxima de 6 KHz y correctamente adquirida mediante un sistema de sonido estéreo con muestras de 16 bits?

A) No más de 1 Mbyte
B) No menos de 5 Mbytes
C) Entre 1 y 2 MBytes
D) Ninguna de las anteriores

TAMAÑO (B) = DURACIÓN (s) * FRECUENCIA (Hz) * TAMAÑO MUESTRA (B) * nº de canales

Tamaño = 10 * 6000 * 2 * 2 (canales, ya que es estéreo) = 240000 B = 0,24 MB.

A ¿Cuál es el espacio requerido para almacenar una colección de 100 documentos representados mediante la presencia o ausencia (1 byte) de sus trigramas (tripletas de tokens) sabiendo que el vocabulario es de 100 tokens?

A) No más de 100 Mbytes
B) No menos de 200 Mbytes
C) Entre 50 y 90 Mbytes
D) Ninguna de las anteriores

D * V * (log2(lmax + 1) / 8) B = 100 * 100^3 * 1 B = 100 MB.

(log2(100 + 1) / 8) = 0,83 -> se redondea a 1.

$\boxed{\text{C}}$ Dado un espacio de representación de d dimensiones se desea aplicar PCA. ¿ Cuál de las siguientes frases es correcta ?

A) Para reducir de dimensionalidad a k dimensiones escogeremos los k mayores eigenvectores (mayor eigenvalor asociado) de la matriz de datos

B) Para reducir de dimensionalidad a k dimensiones escogeremos los k menores eigenvectores (mayor eigenvalor asociado) de la matriz de covarianzas

C) Para reducir de dimensionalidad a k dimensiones escogeremos los k mayores eigenvectores (mayor eigenvalor asociado) de la matriz de covarianzas

D) Para reducir de dimensionalidad a k dimensiones escogeremos los k menores eigenvectores (mayor eigenvalor asociado) de la matriz de datos

$\boxed{\text{B}}$ Dado un problema de clasificación en C clases donde los objetos se representan en un espacio de representación de d dimensiones. Se desea obtener una representación final en un espacio reducido de k dimensiones. Para ello se realizará primero una proyección mediante PCA a d' dimensiones con el fin de evitar singularidades, para posteriormente mediante LDA una proyección final a las k dimensiones. Por lo tanto se debe cumplir que, en general:

A) $d' <= C - 1$ y $k <= d$

B) $k <= min(C - 1, d')$ y $d' <= d$

C) $d' <= min(C - 1, d)$ y $k <= d$

D) $k <= C - 1$ y $d' <= d$

$\boxed{\text{B}}$ Dada la descomposición de la matriz de covarianza de los datos $\mathbf{C}_{3x3}$, en valores y vectores propios: $\lambda_1 = 5.1$ con $\mathbf{w}_1 = (1\ 0\ 0)$, $\lambda_2 = 0.3$ con $\mathbf{w}_2 = (0\ 1\ 0)$, y $\lambda_3 = 2.4$ 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 $\mathbf{w}_2$ y $\mathbf{w}_3$.

B) La proyección PCA de $\mathbb{R}^3$ a $\mathbb{R}^2$ se llevará a cabo con los vectores propios $\mathbf{w}_1$ y $\mathbf{w}_3$.

C) La proyección PCA de $\mathbb{R}^3$ a $\mathbb{R}^1$ se llevará a cabo con el vector propio $\mathbf{w}_2$.

D) La proyección PCA de $\mathbb{R}^3$ a $\mathbb{R}^1$ se llevará a cabo con en vector propio $\mathbf{w}_3$.

$\boxed{\text{A}}$ Cuál de las siguientes afirmaciones respecto a kernels es falsa:

A) Se recomienda emplear funciones kernel solo si el espacio de representación original es linealmente separable

B) Se recomienda emplear funciones kernel cuando el espacio de representación original no es linealmente separable

C) Los kernels modelan el producto escalar en un nuevo espacio de representación

D) La función discriminante es: $g(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i c_i K(\mathbf{x}, \mathbf{x}_i) \ + \ \sum_{i=1}^{n} \alpha_i c_i$

$\boxed{\text{D}}$ Sean $K_1(\mathbf{x}, \mathbf{y})$ y $K_2(\mathbf{x}, \mathbf{y})$ funciones kernel, indica cuál de las siguientes composiciones no es una función kernel:

A) $(c + K_1(\mathbf{x}, \mathbf{y}))^d$ con $d, c > 0$

B) $K_1(\mathbf{x}, \mathbf{y}) + K_2(\mathbf{x}, \mathbf{y})$

C) $exp(K_1(\mathbf{x}, \mathbf{y}))$

D) $c \cdot K_1(\mathbf{x}, \mathbf{y})$

Examen de Teoría de Percepción - Recuperación Primer 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)** Sean un problema de clasificación en dos clases donde las muestras se representan en un espacio vectorial de 2 dimensiones, $\mathbf{x} = (x_1, x_2) \in \mathcal{R}^2$. Se dispone de un clasificador basado en FDL's:

$$g_A(\mathbf{x}) = 2x_1 + x_2 + 1$$
$$g_B(\mathbf{x}) = x_1 - x_2 - 1$$

a) Calcula la frontera de decisión

b) Clasifica la muestra $\mathbf{y} = (2, 0)$

c) Suponiendo que el espacio de representación original es $\mathcal{R}^4$ y las muestras son proyectadas a $\mathcal{R}^2$ mediante la matriz de proyección W:

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

d) Clasifica la muestra $\mathbf{y}' = (1, 0, 1, 1)$

Solución:

a) La recta: $2x_1 + x_2 + 1 = x_1 - x_2 - 1; x_1 + 2x_2 + 2 = 0;$

b) $g_A(\mathbf{y}) = 5; g_B(\mathbf{y}) = 1; \rightarrow$ Clase A

c) Al proyectar $\mathbf{y}'$ nos queda $\mathbf{y} = (2, 0)$ igual que el apartado anterior, clase A.

2. **(2 puntos)** Sea la siguiente función kernel, $K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^2$ y sea el siguiente conjunto de aprendizaje $X = \{(\mathbf{x}_1, +1), (\mathbf{x}_2, +1), (\mathbf{x}_3, +1), (\mathbf{x}_4, -1), (\mathbf{x}_5, -1), (\mathbf{x}_6, -1)\}$ con:

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \mathbf{x}_6 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Se pide:

a) Obtén la matriz kernel K de las muestras de entrenamiento

b) Realiza una iteración (desde $\mathbf{x}_1$ hasta $\mathbf{x}_6$) del algoritmo Kernel Perceptron

c) Clasifica la muestra de test $\mathbf{x} = (-1 \quad 1)$ con el valor de los pesos α obtenido en el apartado anterior

Solución:

a) $K = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 1 & 0 & 4 & 9 \\ 1 & 1 & 4 & 0 & 4 & 4 \\ 1 & 0 & 0 & 9 & 1 & 4 \\ 1 & 4 & 4 & 1 & 9 & 16 \\ 1 & 9 & 4 & 4 & 16 & 36 \end{bmatrix}$

b) $x_1 \rightarrow \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0)$
$x_4 \rightarrow \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0)$
$x_5 \rightarrow \text{error} \Rightarrow \alpha = (1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0)$

Final: $\alpha = (1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0)$

c) $g(\mathbf{x}) = -2 \Rightarrow c(\mathbf{x}) = -1$