Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Abril de 2022

Apellidos: ☐ **Nombre:** ☐

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

Cuestiones (1.5 puntos, 30 minutos, sin apuntes)

B ¿Cuál de las siguientes expresiones equivale a un clasificador de Bayes?

A) $\arg\max_{c\in\mathbb{C}} P(c|\mathbf{x})^{-1}$
B) $\arg\max_{c\in\mathbb{C}} \log P(c,\mathbf{x})$
C) $\arg\min_{c\in\mathbb{C}} \log P(c,\mathbf{x})$
D) $\arg\min_{c\in\mathbb{C}} P(c|\mathbf{x})$

C La técnica del *on-line learning* se caracteriza por:

A) Realizar la selección de las muestras de entrenamiento que aportan más información
B) Emplear únicamente muestras de la realimentación del usuario
C) Usar interpolación para actualizar el modelo
D) Aplicarse sólo a clasificadores bayesianos

D Si se tiene una imagen donde se desea reconstruir con fidelidad tamaños de 1mm, la frecuencia mínima de muestreo debe ser:

A) Entre 1000 y 1250 puntos por metro
B) Entre 1250 y 1500 puntos por metro
C) Entre 1500 y 1750 puntos por metro
D) Mayor a 1750 puntos por metro

B Tras la tokenización y la normalización, ¿cómo queda el texto "Ante esto, podría ir a Madrid."?

A) "Ante esto , podría ir a Madrid ."
B) "ante esto , podria ir a madrid ."
C) "ante esto , podria ir a Madrid ."
D) "ante esto, podria ir a madrid."

$\boxed{\text{A}}$ PCA es una técnica de reducción de dimensión que:

A) Obtiene vectores de proyección ortonormales
B) Minimiza el error de clasificación
C) Minimiza la varianza de los datos proyectados
D) No se basa en un problema de optimización

$\boxed{\text{C}}$ La matriz de covarianzas de los datos proyectados por PCA:

A) Tiene diagonal nula
B) Es la matriz identidad
C) Es una matriz diagonal
D) Es una matriz no simétrica

$\boxed{\text{D}}$ ¿Cuál de los siguientes clasificadores es equivalente al clasificador de un vecino más cercano en distancia L2?

A) $\hat{c}(\mathbf{y}) = \arg\min_c \text{mín}_{\mathbf{x}\in X_c} \frac{1}{L2(\mathbf{x},\mathbf{y})}$
B) $\hat{c}(\mathbf{y}) = \arg\min_c \text{máx}_{\mathbf{x}\in X_c} \frac{1}{L2(\mathbf{x},\mathbf{y})}$
C) $\hat{c}(\mathbf{y}) = \arg\max_c \text{mín}_{\mathbf{x}\in X_c} \frac{1}{L2(\mathbf{x},\mathbf{y})}$
D) $\hat{c}(\mathbf{y}) = \arg\max_c \text{máx}_{\mathbf{x}\in X_c} \frac{1}{L2(\mathbf{x},\mathbf{y})}$

$\boxed{\text{C}}$ Si se tiene un error de Bayes de $P = 0.25$, ¿qué error se puede esperar con un número asintóticamente grande de prototipos y vecinos en un clasificador k-NN?

A) 0.5 en cualquier caso
B) 0.25, siempre que k crezca más rápido que el número de prototipos
C) 0.25, siempre que k crezca más lento que el número de prototipos
D) Dependerá del número de clases

Examen de Teoría de Percepción - Primer Parcial

ETSINF, Universitat Politècnica de València, Abril de 2022

Apellidos: ☐ **Nombre:** ☐

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

Problemas (2 puntos, 90 minutos, con apuntes)

1. (**0.5 puntos**) Calcula el espacio en memoria de las siguientes representaciones:

   a) Representación global por histograma de una imagen a 1024 niveles de gris con resolución 512×256 píxeles (**0.05 puntos**)

   b) Representación local de una imagen de 2048×1024 píxeles, usando ventanas de 5×15 píxeles y una rejilla de desplazamiento horizontal de 4 y vertical de 4 sobre una imagen de 512 niveles de gris, usando representación directa de cada ventana (**0.2 puntos**)

   c) Señal de audio de 5 canales de 20 minutos de duración, con ancho de banda 48KHz y 32 bits (**0.1 puntos**)

   d) Colección de 2000 documentos de 1000 palabras máximo cada uno, con un vocabulario de 2000 palabras, representado por *term frequency* de 2-grama (**0.15 puntos**)

   Solución:

   a) 3 Kbytes

   b) 18.49 Mbytes

   c) 2197.27 Mbytes

   d) 15258.79 Mbytes

2. **(0.8 puntos)** Se tiene el siguiente conjunto de datos vectoriales de 3 dimensiones ($\mathbf{x} \in \mathbb{R}^3$) con sus correspondientes etiquetas de clase:

n	1	2	3	4	5	6
x_{n1}	2	-1	2	-2	-2	1
x_{n2}	1	1	-2	-1	2	-1
x_{n3}	-1	-1	2	1	-2	1
c_n	A	B	C	A	B	C

   Se pide:

   a) Calcular una matriz de proyección a dos dimensiones ($\mathbb{R}^2$) mediante PCA, indicando todos los pasos necesarios (**0.5 puntos**)

   b) Aplicar dicha proyección y discernir si los datos proyectados serían lineamente separables. En ese caso, proponer un clasificador basado en funciones discriminantes lineales que clasifiquen correctamente las muestras proyectadas (**0.3 puntos**)

   Solución:

   a) La media de los datos es $\mathbf{x}_m = (0 \quad 0 \quad 0)^t$ y la matriz de covarianzas de esos datos es:

$$\Sigma_X = \begin{pmatrix} 3 & -1 & 1 \\ -1 & 2 & -2 \\ 1 & -2 & 2 \end{pmatrix}$$

   Los valores y vectores propios de la matriz de covarianzas son:

	$\mathbf{w}_1$	$\mathbf{w}_2$	$\mathbf{w}_3$
	$-\frac{2}{\sqrt{6}}$	$-\frac{1}{\sqrt{3}}$	0
	$-\frac{1}{\sqrt{6}}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{2}}$
	$\frac{1}{\sqrt{6}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{2}}$
λ	2	5	0

Por tanto, la matriz de proyección a $\mathbb{R}^2$ se realiza mediante los vectores $\mathbf{w}_2$ y $\mathbf{w}_1$:

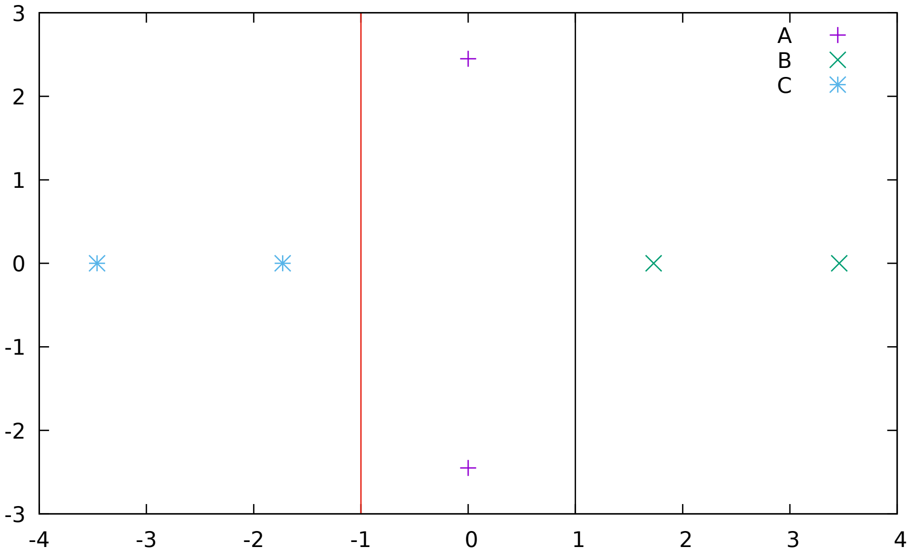
$$W = \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

b) El resultado de la proyección es:

$\mathbf{x}'_1$	$\mathbf{x}'_2$	$\mathbf{x}'_3$	$\mathbf{x}'_4$	$\mathbf{x}'_5$	$\mathbf{x}'_6$
0	$\sqrt{3}$	$-2\sqrt{3}$	0	$2\sqrt{3}$	$-\sqrt{3}$
$-\sqrt{6}$	0	0	$\sqrt{6}$	0	0
A	B	C	A	B	C

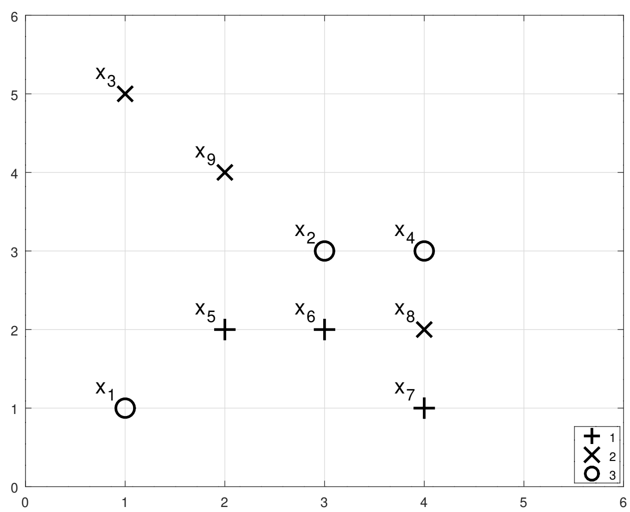
Son linealmente separables (ver la representación gráfica) mediante las fronteras verticales $-x_1 = 1$ y $1 = x_1$, que han sido obtenidas como resultado de igualar las funciones discriminantes $g_C(\mathbf{x}) = g_A(\mathbf{x})$ y $g_A(\mathbf{x}) = g_B(\mathbf{x})$, respectivamente. Es decir:

$$g_A(\mathbf{x}) = 1 \quad g_B(\mathbf{x}) = x_1 \quad g_C(\mathbf{x}) = -x_1$$



3. **(0.7 puntos)** Se tiene el siguiente conjunto de datos, cuya representación gráfica se ve en la parte derecha:

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



Se pide:

a) Aplica el algoritmo de Wilson con 1-NN en distancia Euclídea, con recorrido por índices ascendentes. En caso de empate por distancia, desempata clasificando por el prototipo de menor distancia en x_{n2} y si persiste el empate al de menor valor de x_{n1} entre los empatados (el que esté más a la izquierda en la gráfica) (**0.4 puntos**)

b) Una vez aplicado el algoritmo de Wilson, aplica el algoritmo de Hart con 1-NN en distancia Euclídea, con recorrido por índices ascendentes. Utiliza los mismo criterios en caso de empate que en el apartado anterior (**0.3 puntos**)

Solución:

a) Tras aplicar Wilson, el conjunto resultante de prototipos es $\{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4, \mathbf{x}_5, \mathbf{x}_6, \mathbf{x}_9\}$.

b) Una vez aplicado Wilson, aplicamos Hart, y obtenemos los conjuntos $S = \{\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_5\}$ y $G = \{\mathbf{x}_4, \mathbf{x}_6, \mathbf{x}_9\}$.