

Apellidos:  Nombre:

Profesor: ☒ Jorge Civera ☐ Carlos Martínez

Cuestiones (0.25 puntos, 15 minutos, con apuntes)

☐ A Dado el siguiente conjunto de datos en  $\mathbb{R}^2$ :

n	1	2	3	4	5	6	7	8	9	10
$x_1$	1	3	1	-3	0	2	1	-2	-2	1
$x_2$	-2	1	3	0	-2	1	2	1	0	1
$c_n$	D	B	A	C	D	B	A	C	C	A

Indicar la clasificación por vecino más cercano de la muestra  $\mathbf{y} = (0, 0)^t$  usando la distancia  $L_\infty/L_0$ .

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

☐ C En un clasificador  $k$ -NN siendo  $k$  un número impar, ¿qué tipo de empates se pueden producir?

- A) Sólo por distancia entre una o más clases para definir los  $k$  vecinos
- B) Sólo por número de vecinos de la misma clase entre los  $k$  vecinos
- C) Tanto por distancia como por número de vecinos de la misma clase
- D) Ni por distancia ni por número de vecinos de la misma clase

☐ D Dado el conjunto de prototipos ya editado  $X = \{((0, 0), A), ((1, 0), A), ((0, 2), B), ((1, 2), B)\}$ , ¿cuántos conjuntos diferentes  $S$  se pueden obtener tras aplicar el algoritmo de condensado de Hart en función del recorrido que se haga de los prototipos?

- A) 1
- B) 2
- C) 3
- D) 4

Apellidos:  Nombre:

Profesor: ☐ Jorge Civera ☒ Carlos Martínez

Cuestiones (0.25 puntos, 15 minutos, con apuntes)

☐ D Sea la función  $d : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$  tal que  $d(\mathbf{x}, \mathbf{y}) = \log \sum_{i=1}^D |x_i - y_i|$

- A) Cumple la propiedad nula:  $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$
- B) Cumple la propiedad positiva:  $d(\mathbf{x}, \mathbf{y}) > 0$
- C) Cumple la propiedad positiva y la propiedad nula
- D) No cumple ni la propiedad nula, ni la propiedad positiva

☐ A Si se tiene una tarea en la que el clasificador de Bayes obtiene un óptimo de error de 0.1, ¿qué afirmación sería correcta sobre un clasificador  $k$ -NN sobre esa tarea suponiendo que hay disponibles  $N$  prototipos?

- A) Obtendrá un error mayor o igual a 0.1
- B) Obtendrá un error máximo que dependerá del número de clases de la tarea
- C) Obtendrá un error inferior a 0.1 siempre que  $N \rightarrow \infty$ ,  $k \rightarrow \infty$  y  $\frac{k}{N} \rightarrow 0$
- D) Obtendrá un error mínimo de 0.2

☐ A En la distancia euclídea ponderada, los pesos utilizados son función de la varianza de los datos estimada para todos los datos, en una clase o en el entorno del prototipo a clasificar, ¿cuál es el objetivo de esta ponderación?

- A) A mayor varianza en una dimensión, menor contribución de esta dimensión a la distancia total
- B) A mayor varianza en una dimensión, mayor contribución de esta dimensión a la distancia total
- C) A menor varianza en una dimensión, menor contribución de esta dimensión a la distancia total
- D) Depende de la estimación que se haga de la varianza