Sistemas Inteligentes
Cuestiones y ejercicios del bloque 2, tema 2
Aprendizaje de funciones discriminantes: Perceptrón

Escola Tècnica Superior d'Informàtica
Dep. de Sistemes Informàtics i Computació
Universitat Politècnica de València

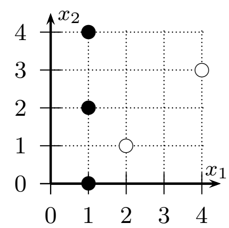
17 de noviembre de 2024

1. Cuestiones

1 ☐ El algoritmo Perceptrón es una técnica de aprendizaje...

A) supervisado de clasificadores lineales.
B) supervisado de clasificadores no-lineales.
C) no-supervisado de clasificadores lineales.
D) no-supervisado de clasificadores no-lineales.

2 ☐ En la figura de la derecha se representan cinco muestras de aprendizaje bidimensionales de 2 clases: ○ y ●. Se desea contruir un clasificador lineal para cualquier $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, como sigue:



$$c(\mathbf{x}) = \begin{cases} \circ & \text{si } \mathbf{w}^t\,\mathbf{x} > 0 \\ \bullet & \text{si } \mathbf{w}^t\,\mathbf{x} \leq 0 \end{cases} \quad \text{donde} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \text{ es un vector de pesos a escoger}$$

Si nuestro criterio de aprendizaje es la minimización del número de errores de clasificación (sobre las muestras de aprendizaje), elegiremos...

A) $\mathbf{w} = (1, 0)^t$
B) $\mathbf{w} = (1, 1)^t$
C) $\mathbf{w} = (1, -1)^t$
D) Ninguna de las anteriores, ya que hay otros vectores de pesos que producen menos errors sobre las muestras dadas.

3 ☐ En el algoritmo Perceptrón:

A) principalmente existen dos parámetros: número de clases y número de prototipos.
B) la constante de aprendizaje tiene que ser lo más alta posible, para que aprenda más.
C) el margen tiene que ser cero cuando las clases no son linealmente separables.
D) principalmente existen dos parámetros: la constante de aprendizaje α y el margen b.

4 ☐ En el algoritmo Perceptrón:

A) principalmente existen dos parámetros: número de clases y número de prototipos.
B) el uso del margen b permite encontrar soluciones adecuadas cuando el problema no es linealmente separable.
C) el valor del margen b depende del valor de la constante de aprendizaje α empleado.
D) principalmente existen dos parámetros: la constante de aprendizaje α y el número de iteraciones.

5 ☐ El parámetro del algoritmo Perceptrón que denominamos *margen*, b, es un valor real que, suponiendo que sea positivo (como suele ser), restringe el conjunto de soluciones a las que puede converger el algoritmo. Concretamente, dadas N muestras de entrenamiento $(\mathbf{x}_1, c_1), \ldots, (\mathbf{x}_N, c_N)$ de C clases, el algoritmo Perceptrón tratará de hallar funciones discriminantes lineales $g_1(\cdot), \ldots, g_C(\cdot)$ tales que, para todo $n = 1, \ldots, N$:

A) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n)$ para toda clase $c \neq c_n$
B) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n) + b$ para toda clase $c \neq c_n$
C) $g_{c_n}(\mathbf{x}_n) > g_c(\mathbf{x}_n) - b$ para toda clase $c \neq c_n$
D) Ninguna de las anteriores

6 □ Sea un problema de clasificación en C clases $C \in \{1, 2, \ldots, C\}$, en el que los objetos están representados mediante puntos en un espacio vectorial de D dimensiones, $\mathbf{x} \in \mathbb{R}^D$. Suponiendo que un $\mathbf{x}$ dado pertenece a la clase 1, el algoritmo Perceptrón:

A) Modifica el discriminante lineal $g_1(\mathbf{x})$ en todo caso.
B) Modifica el discriminante lineal $g_1(\mathbf{x})$ si existe $c \neq 1, g_c(\mathbf{x}) > g_1(\mathbf{x})$.
C) Modifica el discriminante lineal $g_c(\mathbf{x})$ si $g_c(\mathbf{x}) < g_1(\mathbf{x})$ con $c \neq 1$.
D) Modifica el discriminante lineal $g_1(\mathbf{x})$ sólo si $g_c(\mathbf{x}) > g_1(\mathbf{x})$ para todo $c \neq 1$.

7 □ En un problema de clasificación en dos clases se tienen los siguientes puntos en dos dimensiones: $\mathbf{x}_1 = (1,1)^t, \mathbf{x}_2 = (2,2)^t, \mathbf{x}_3 = (2,0)^t$; $\mathbf{x}_1$ y $\mathbf{x}_2$ pertenencen a la clase A y $\mathbf{x}_3$ a la clase B. Teniendo en cuenta que se emplea un clasificador basado en funciones discriminantes lineales con vectores de pesos $\mathbf{w}_A$ y $\mathbf{w}_B$ asociados a las clases A y B respectivamente, indica cuál de las siguientes afirmaciones es *falsa*:

A) Se puede encontrar una función discriminante lineal que clasifique $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=2/3.
B) Los pesos $\mathbf{w}_A = (1, -1, 1)^t$ y $\mathbf{w}_B = (1, 2, -4)^t$ clasifican $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ sin error.
C) Los pesos $\mathbf{w}_A = (1, -1, 1)^t$ y $\mathbf{w}_B = (1, 2, -4)^t$ clasifican $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=1/3.
D) Se puede encontrar una función discriminante lineal que clasifique $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ con error=1/3.

8 □ Sea un problema de clasificación en tres clases $\{A, B, C\}$. El espacio de representación de los objetos es bidimensional, $\mathbb{R}^2$. Se propone emplear un clasificador basado en funciones discriminantes lineales con los siguientes vectores de pesos para cada clase, $\mathbf{w}_A = (1, 1, 0)^t$, $\mathbf{w}_B = (-1, 1, -1)^t$ y $\mathbf{w}_C = (1, -2, 2)^t$. ¿Cuál es la clasificación de los objetos $\mathbf{x}_1 = (1, 1)^t$ y $\mathbf{x}_2 = (0, -1)^t$?

A) $c(\mathbf{x}_1) = B \quad c(\mathbf{x}_2) = C$
B) $c(\mathbf{x}_1) = A \quad c(\mathbf{x}_2) = B$
C) $c(\mathbf{x}_1) = B \quad c(\mathbf{x}_2) = A$
D) $c(\mathbf{x}_1) = A \quad c(\mathbf{x}_2) = A$

9 □ (Examen de SIN del 18 de enero de 2013)
Sean $g_1(\mathbf{y}) = y_1^2 + 2\,y_2^2$ y $g_2(\mathbf{y}) = 2\,y_1^2 + y_2^2$ dos funciones discriminantes para las clases 1 y 2, respectivamente. La frontera de decisiónentre estas clases es:

A) Una parábola.
B) Hiperesférica.
C) Viene dada por la ecuación $y_1^2 + y_2^2 = 0$.
D) Una recta.

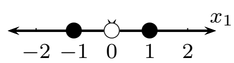
10 □ (Examen de SIN del 30 de enero de 2013)
Para un problema de clasificación de dos clases en $\Re^2$ se han construido tres clasificadores distintos. Uno está formado por las dos funciones discriminantes lineales siguientes: $g_1(y) = 3 + 4\,y_1 - 2\,y_2$ y $g_2(y) = -3 + 1.5\,y_1 + 5\,y_2$. El segundo clasificador por $g_1'(y) = 6 + 8\,y_1 - 4\,y_2$ y $g_2'(y) = -6 + 3\,y_1 + 10\,y_2$. El tercero por $g_1''(y) = -6 - 8\,y_1 + 4\,y_2$ y $g_2''(y) = 6 - 3\,y_1 - 10\,y_2$. ¿Los tres clasicadores son equivalentes? es decir ¿definen las mismas fronteras de decisión?

A) (g_1, g_2) y (g_1', g_2') son equivalentes.
B) Los tres son equivalentes.
C) (g_1, g_2) y (g_1'', g_2'') son equivalentes.
D) (g_1', g_2') y (g_1'', g_2'') son equivalentes.

11 □ (Examen de SIN del 30 de enero de 2013)
El algoritmo Perceptrón es una técnica de aprendizaje...

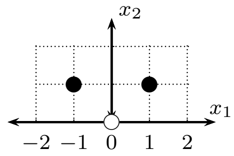
A) supervisado de clasificadores lineales.
B) supervisado de clasificadores cuadráticos.
C) no-supervisado de clasificadores lineales.
D) no-supervisado de clasificadores cuadráticos.

12 □ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 4)
En la figura de la derecha se representan tres muestras de aprendizaje unidimensionales de 2 clases: ○ y ●. ¿Cuál será el número de errores de clasificación cometidos por un clasificador lineal de mínimo error?



A) 0
B) 1
C) 2
D) 3

13 ☐ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 5)
Supongamos que en el ejercicio anterior añadimos un nueva característica x_2 que se define como $x_2 = x_1^2$. De esta forma las tres muestras de aprendizaje pasan a ser bidimensionales como se observa en la figura de la derecha. En este caso, ¿cuál será el número de errores de clasificación cometidos por un clasificador lineal de mínimo error?



A) 0
B) 1
C) 2
D) 3

14 ☐ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 6)
Sea un problema de clasificación en 2 clases, $c = 1, 2$, para objetos representados mediante vectores de características bidimensionales. Se tienen 2 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, y $\mathbf{x}_2 = (1, 1)^t$ de $c_2 = 2$. Asimismo, se tiene un clasificador lineal definido por los vectores de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (1, -1, -1)^t$ y $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-1, 1, 1)^t$. Si aplicamos una iteración del algoritmo Perceptrón a partir de estos vectores de pesos, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, entonces:

A) No se modificará ningún vector de pesos.
B) Se modificará el vector de pesos de la clase 1.
C) Se modificará el vector de pesos de la clase 2.
D) Se modificarán los vectores de pesos de ambas clases.

15 ☐ (Examen de SIN del 15 de enero de 2014; examen del bloque 2; cuestión 7)
El algoritmo Perceptrón está gobernado por dos parámetros que denominamos *velocidad de aprendizaje*, α, y *margen*, b, siendo ambos valores reales. En caso de que no supiéramos si las muestras de aprendizaje son linealmente separables, ¿qué valores de los parámetros α y b proporcionan mayores garantías de obtener fronteras de decisión de mejor calidad?

A) $\alpha = 0.1$ y $b = 0.0$.
B) $\alpha = 0.0$ y $b = 0.0$.
C) $\alpha = 0.1$ y $b = 1.0$.
D) $\alpha = 0.0$ y $b = 1.0$.

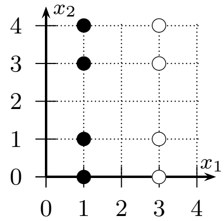
16 ☐ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 1)
En un experimento de clasificación con 300 datos de *test* se han observado 15 errores. Con una confianza del 95 %, podemos afirmar que la verdadera probabilidad de error es:

A) $P(\text{error}) = 5\,\% \pm 0.3\,\%$
B) $P(\text{error}) = 0.05 \pm 0.3$
C) $P(\text{error}) = 0.05$, exactamente
D) $P(\text{error}) = 0.05 \pm 0.03$

17 ☐ (Examen de SIN del 28 de enero de 2014; examen final; cuestión 3)
Sea un problema de clasificación en 2 clases, $c = A, B$, para objetos representados mediante vectores de características bidimensionales. Como resultado de la aplicación del algoritmo Perceptrón sobre un conjunto de entrenamiento, se han obtenido los vectores de pesos $\mathbf{w}_A = (1, 1, 0)^t$ y $\mathbf{w}_B = (-1, 0, 1)^t$. ¿En qué clases se clasifican $\mathbf{x}_1 = (-1, 0)^t$ y $\mathbf{x}_2 = (0, 3)^t$?

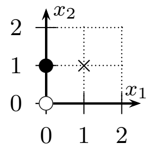
A) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = A$.
B) $\hat{c}(\mathbf{x}_1) = A$ y $\hat{c}(\mathbf{x}_2) = B$.
C) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = A$.
D) $\hat{c}(\mathbf{x}_1) = B$ y $\hat{c}(\mathbf{x}_2) = B$.

18 ☐ En la figura de la derecha se representan cuatro muestras de aprendizaje bidimensionales de 2 clases: ○ y ●. Si nuestro criterio de aprendizaje es la minimización del número de errores de clasificación (sobre las muestras de aprendizaje), elegiremos como vector de pesos de cada una de las clases…



A) $\mathbf{a}_\circ = (3, 1, 1)^t$ y $\mathbf{a}_\bullet = (1, 2, 1)^t$
B) $\mathbf{a}_\circ = (1, 1, 2)^t$ y $\mathbf{a}_\bullet = (3, 1, 1)^t$
C) $\mathbf{a}_\circ = (3, 1, 1)^t$ y $\mathbf{a}_\bullet = (1, 1, 2)^t$
D) $\mathbf{a}_\circ = (1, 2, 1)^t$ y $\mathbf{a}_\bullet = (3, 1, 1)^t$

19 En la figura de la derecha se representan tres muestras de aprendizaje bidimensionales de 3 clases: $\circ$, $\bullet$ y $\times$. Dados el conjunto de pesos $\mathbf{a}_\circ = (-2,-1,-3)^t$, $\mathbf{a}_\bullet = (-1,-3,1)^t$ y $\mathbf{a}_\times = (-3,3,-1)^t$, ¿cuántos errores de clasificación se producen sobre las muestras de aprendizaje?



A) 0
B) 1
C) 2
D) 3

20 Si aplicamos una iteración del algoritmo Perceptrón con factor de aprendizaje $\alpha = 1.0$ y margen $b = 0.0$ a partir del conjunto de pesos y muestras de aprendizaje de la cuestión anterior, ¿cuántos errores de clasificación se producen sobre las muestras de aprendizaje con el nuevo conjunto de pesos?

A) 0

B) 1

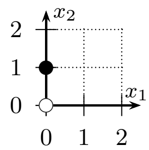
C) 2

D) 3

21 Dado un clasificador lineal de 2 clases $\circ$ y $\bullet$ definido por su conjunto de pesos $\mathbf{a}_\circ = (0,-1,1)^t$ y $\mathbf{a}_\bullet = (0,1,-1)^t$, ¿Qué conjunto de pesos de los siguientes no define un clasificador equivalente al dado?

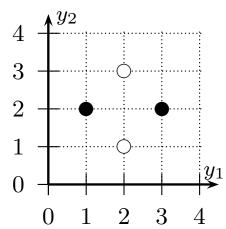
A) $\mathbf{a}_\circ = (1,-1,1)^t$ y $\mathbf{a}_\bullet = (1,1,-1)^t$
B) $\mathbf{a}_\circ = (-1,-2,2)^t$ y $\mathbf{a}_\bullet = (-1,2,-2)^t$
C) $\mathbf{a}_\circ = (0,2,-2)^t$ y $\mathbf{a}_\bullet = (0,-2,2)^t$
D) $\mathbf{a}_\circ = (0,-2,2)^t$ y $\mathbf{a}_\bullet = (0,2,-2)^t$

22 En la figura de la derecha se representan dos muestras de aprendizaje bidimensionales de 2 clases: (x_1,$\circ$) y (x_2,$\bullet$). Dados el conjunto de pesos $\mathbf{a}_\circ = (0,1,-2)^t$ y $\mathbf{a}_\bullet = (0,0,1)^t$, si aplicamos una iteración del algoritmo Perceptrón con factor de aprendizaje $\alpha = 1.0$ y margen $b = 0.5$ a partir del conjunto de pesos y muestras de aprendizaje dadas, ¿cuántos errores de clasificación se producen sobre las muestras de aprendizaje con el nuevo conjunto de pesos?



A) 0
B) 1
C) 2
D) 3

23 En la figura de la derecha se representan cuatro muestras de aprendizaje bidimensionales de 2 clases: $\circ$ y $\bullet$. A estas muestras se les aplica el algoritmo Perceptrón con pesos iniciales $\mathbf{a}_\circ = (0,1,0)^t$ y $\mathbf{a}_\bullet = (0,0,1)^t$, una constante de aprendizaje $\alpha > 0$ y un margen b. Indica cuál de las siguientes afirmaciones es correcta:



A) El algoritmo convergerá para algún $b > 0$.

B) El algoritmo solo puede converger si $b \leq 0$.

C) Si $b > 0$, no hay convergencia, pero se puede ajustar el valor de α tal que, tras un número finito de iteraciones, se obtengan buenas soluciones (con 25 % de error de resustitución).

D) El algoritmo no es aplicable a eastas muestras porque no son linealmente separables.

24 ¿Cuál sería el número mínimo de errores de un clasificador lineal en el conjunto de muestras de la cuestión anterior?

A) 0
B) 1
C) 2
D) 3

25 Dado un clasificador lineal de 2 clases $\circ$ y $\bullet$ definido por su conjunto de pesos $\mathbf{a}_\circ = (3,1,1)^t$ y $\mathbf{a}_\bullet = (1,2,1)^t$ (en notación homogénea, cuya primera componente es el término independiente de la función lineal correspondiente). ¿Cuál de las siguientes afirmaciones es correcta?

A) Como hay dos vectores de pesos y el espacio de representación es bi-dimensional, tendremos 4 regiones de decisión.

B) Los vectores de pesos $\mathbf{a}_\circ = (2, -2, -2)^t$ y $\mathbf{a}_\bullet = (-2, 0, -2)^t$ determinan la misma frontera de decisión que la del clasificador dado.

C) Un clasificador equivalente al dado es el definido por $\mathbf{a}_\circ = (1, 2, 1)^t$ y $\mathbf{a}_\bullet = (3, 1, 1)^t$.

D) Como los vectores de pesos son de tres dimensiones, la frontera viene dada por la ecuación de un plano en $\mathbb{R}^3$.

26 ☐ Para un problema de clasificación de dos clases en $\Re^2$ se han construido tres clasificadores distintos. Uno está formado por las dos funciones discriminantes lineales siguientes: $g_1(y) = 2\,y_1 + y_2 + 3$ y $g_2(y) = y_1 + 2$. El segundo clasificador por $g_1'(y) = -2\,y_1 + y_2 - 1$ y $g_2'(y) = -y_1 + 2\,y_2$. El tercero por $g_1''(y) = -2\,y_1 - y_2 - 3$ y $g_2''(y) = -y_1 - 2$. ¿Cuál de las siguientes afirmaciones es cierta?

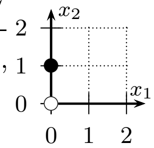
A) (g_1, g_2) y (g_1', g_2') son equivalentes, pero (g_1, g_2) y (g_1'', g_2'') no lo son.

B) (g_1, g_2) y (g_1', g_2') no son equivalentes, pero (g_1, g_2) y (g_1'', g_2'') lo son.

C) (g_1, g_2) y (g_1', g_2') no son equivalentes, pero (g_1', g_2') y (g_1'', g_2'') lo son.

D) Los tres no son equivalentes entre sí.

27 ☐ En la figura de la derecha se representan dos muestras de aprendizaje bidimensionales de 2 clases: $(x_1,\circ)$ y $(x_2,\bullet)$. Dados el conjunto de pesos $\mathbf{a}_\circ = (0, 1, -2)^t$ y $\mathbf{a}_\bullet = (0, 0, 1)^t$, si aplicamos el algoritmo Perceptrón procesando únicamente la muestra x_1, obtenemos un nuevo conjunto de pesos $\mathbf{a}_\circ = (1, 1, -2)^t$ y $\mathbf{a}_\bullet = (-1, 0, 1)^t$, ¿qué valor tienen el factor de aprendizaje α y el margen b?



A) $\alpha = 1.0$ y $b = 0.0$

B) $\alpha = -1.0$ y $b = 0.5$

C) $\alpha = 1.0$ y $b = 0.5$

D) No es posible determinar los valores de α y b

28 ☐ Para un problema de clasificación de dos clases en $\mathbb{R}^2$ tenemos un clasificador compuesto por dos funciones discriminantes lineales cuyos vectores de pesos en notación homogénea son $\mathbf{a}_\circ = (-1, 1, 2)^t$ y $\mathbf{a}_\bullet = (1, 1, 1)^t$. Indica cuáles son las regiones de decisión definidas por el clasificador anterior.

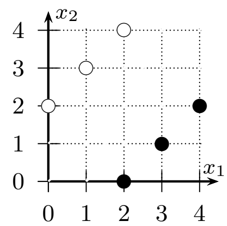
A) $R_\circ = \{\mathbf{x} \in \mathbb{R}^2 : x_2 > 2\}$ y $R_\bullet = \{\mathbf{x} \in \mathbb{R}^2 : x_2 < 2\}$

B) $R_\circ = \{\mathbf{x} \in \mathbb{R}^2 : x_1 > 2\}$ y $R_\bullet = \{\mathbf{x} \in \mathbb{R}^2 : x_1 < 2\}$

C) $R_\circ = \{\mathbf{x} \in \mathbb{R}^2 : x_1 < 2\}$ y $R_\bullet = \{\mathbf{x} \in \mathbb{R}^2 : x_1 > 2\}$

D) $R_\circ = \{\mathbf{x} \in \mathbb{R}^2 : x_2 < 2\}$ y $R_\bullet = \{\mathbf{x} \in \mathbb{R}^2 : x_2 > 2\}$

29 ☐ En la figura de la derecha se representan 6 muestras de aprendizaje bidimensionales de 2 clases. Tras aplicar el algoritmo Perceptrón con diferentes valores del parámetro b, se obtienen los siguientes clasificadores caracterizados por sus vectores de pesos, ¿cuál de ellos proporciona una frontera de decisión con mayor holgura ("centradas"), y por tanto de menor error esperado?



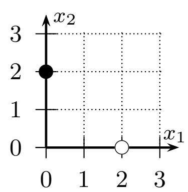
A) $\mathbf{a}_\circ = (-1, 1, 2)^t$ y $\mathbf{a}_\bullet = (\ 0, 2, 1)^t$

B) $\mathbf{a}_\circ = (\ 1, 1, 2)^t$ y $\mathbf{a}_\bullet = (\ 1, 2, 1)^t$

C) $\mathbf{a}_\circ = (\ 1, 1, 2)^t$ y $\mathbf{a}_\bullet = (\ 0, 2, 1)^t$

D) $\mathbf{a}_\circ = (\ 1, 1, 1)^t$ y $\mathbf{a}_\bullet = (-1, 3, 0)^t$

30 ☐ En la figura de la derecha se representan dos muestras de aprendizaje bidimensionales de 2 clases: $(\mathbf{x}_1,\circ)$ y $(\mathbf{x}_2,\bullet)$. Dados los pesos $\mathbf{a}_\circ = (0, 1, 0)^t$ y $\mathbf{a}_\bullet = (1, 0, 0)^t$, si aplicamos una iteración del algoritmo Perceptrón procesando únicamente la muestra $\mathbf{x}_1$, ¿cuál es el valor mínimo del margen b con el que se actualizan los vectores de pesos?



A) $b = 0.5$

B) $b = 1.0$

C) $b = 1.5$

D) Ninguno de los anteriores

31 ☐ Se tiene un problema de clasificación para el cual se ha aprendido un clasificador. El intervalo de confianza al 95 % para la probabilidad de error de dicho clasificador se ha estimado empíricamente, a partir de un cierto conjunto de muestras de test. Indica cuál de las siguientes opciones permitiría reducir el tamaño del intervalo estimado:

A) Reducir significativamente el conjunto de test.
B) Mantener el conjunto de test y re-entrenar el clasificador con el algoritmo C-medias de Duda y Hart.
C) Mantener el conjunto de test y re-entrenar el clasificador con el algoritmo C-medias convencional ("popular").
D) Aumentar significativamente el conjunto de test.

32 ☐ Sean un problema de clasificación de tres clases en $\mathbb{R}^2$ y un clasificador definido por tres funciones discriminantes lineales: $g_1(\mathbf{x})=x_1$, $g_2(\mathbf{x})=x_2$ y $g_3(\mathbf{x})=-x_2$. Indica cuál de las afirmaciones sobre dicho clasificador *no* es correcta.

A) Define tres fronteras de decisión que intersectan en el origen de coordenadas.

B) La región de decisión de la clase 1 se define como $R_1=\{\mathbf{x}\in\mathbb{R}^2 : x_1>0 \wedge x_1>|x_2|\}$.

C) En la región de decisión R_2, x_2 es menor que cero y en R_3, x_2 es mayor que cero.

D) En la región de decisión R_2, x_2 es mayor que cero y en R_3, x_2 es menor que cero.

33 ☐ Sea $\hat{p}$ la probabilidad de error de un clasificador estimada a partir de un conjunto de test cuya talla es N y sea $I = [\hat{p} \pm \epsilon]$ el intervalo de confianza de esta estimación. Indica la respuesta *correcta*.

A) Si $N = 160$ y el clasificador produce al menos un error, ϵ será menor que 1 %.
B) Si $N > 150$ y se obtiene $\hat{p} = 0.1$, ϵ será menor que 5 %.
C) Si N_e es el número de errores del clasificador, entonces $\hat{p} = N/N_e$ y ϵ es inversamente proporcional a N.
D) No es posible determinar ϵ si $\hat{p} = 0$.

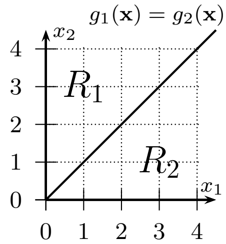
34 ☐ Sea un clasificador en 3 clases para $\mathbf{x} = (x_1, x_2)^t \in [0, 1]^2$ con las distribuciones de probabilidad dadas a la derecha. ¿Cuál es la probabilidad de error p_e del clasificador?

A) $p_e < 0.35$
B) $0.35 \leq p_e < 0.45$
C) $0.45 \leq p_e < 0.65$
D) $0.65 \leq p_e$

x_1	x_2	$p(c = 1\|\mathbf{x})$	$p(c = 2\|\mathbf{x})$	$p(c = 3\|\mathbf{x})$	$p(\mathbf{x})$
0	0	1.0	0.0	0.0	0.1
0	1	0.01	0.01	0.98	0.2
1	0	0.25	0.5	0.25	0.3
1	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0.4

35 ☐ Sea un problema de clasificación en cuatro clases de objetos en $\mathbb{R}^3$. Se tiene un clasificador de funciones discriminantes lineales con vectores de pesos (en notación homogénea): $\mathbf{w}_1 = (-2, 1, 2, 0)^t$, $\mathbf{w}_2 = (0, 2, 2, 0)^t$, $\mathbf{w}_3 = (1, 1, 1, 0)^t$ y $\mathbf{w}_4 = (3, 0, 0, 1)^t$. Indica a qué clase se asignará el objeto $\mathbf{x} = (1, 2, 2)^t$ (*no* en notación homógenea).
A) 1.
B) 2.
C) 3.
D) 4.

36 ☐ En la figura se representan frontera y regiones de decisión de un clasificador binario. ¿Cuál de los siguientes pares de vectores de pesos corresponde al clasificador de la figura?
A) $\mathbf{w}_1 = (-1, -1, -2)^t$ y $\mathbf{w}_2 = (-1, -2, -1)^t$
B) $\mathbf{w}_1 = (\ 1, -1, -2)^t$ y $\mathbf{w}_2 = (\ 0, -2, -1)^t$
C) $\mathbf{w}_1 = (\ 1,\ 1,\ 2)^t$ y $\mathbf{w}_2 = (\ 1,\ 2,\ 1)^t$
D) $\mathbf{w}_1 = (-1,\ 1,\ 2)^t$ y $\mathbf{w}_2 = (\ 0,\ 2,\ 1)^t$

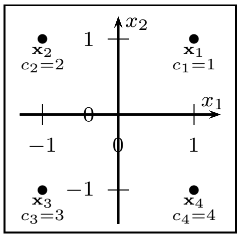


37 ☐ Sea un problema de clasificación en 3 clases, $c = 1, 2, 3$, para objetos representados mediante vectores de características bidimensionales. Se tienen 3 muestras de entrenamiento representadas en notación homogénea: $\mathbf{x}_1 = (1, 1, 2)^t$ de la clase $c_1 = 1$, $\mathbf{x}_2 = (1, 2, 3)^t$ de la clase $c_2 = 2$ y $\mathbf{x}_3 = (1, 3, 1)^t$ de la clase $c_3 = 3$. Asimismo, se tiene un clasificador lineal definido por los vectores de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (2, -8, 0)^t$, $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-5, -2, -1)^t$ y $\mathbf{w}_3 = (w_{30}, w_{31}, w_{32}) = (-2, 1, -10)^t$. Si aplicamos una iteración del algoritmo Perceptrón a partir de estos vectores de pesos, con factor de aprendizaje $\alpha = 1$ y margen $b = 1.5$, entonces:
A) Se modificarán los vectores de pesos $\mathbf{w}_1$ y $\mathbf{w}_2$.
B) Se modificarán los vectores de pesos $\mathbf{w}_1$ y $\mathbf{w}_3$.
C) Se modificarán los vectores de pesos $\mathbf{w}_2$ y $\mathbf{w}_3$.
D) No se modificará ningún vector de pesos.

38 ☐ Se tiene un problema de clasificación en 3 clases, $c = 1, 2, 3$, para objetos representados mediante vectores de 2 características reales, $\mathbf{x} = (x_1, x_2)^t \in \mathbb{R}^2$. Considérese un clasificador lineal de vectores de pesos (en notación homogénea): $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12})^t = (2, 0, 0)^t$, $\mathbf{w}_2 = (0, 1, 1)^t$ y $\mathbf{w}_3 = (0, 1, -1)^t$. La región de decisión de la clase 1 correspondiente a este clasificador es:

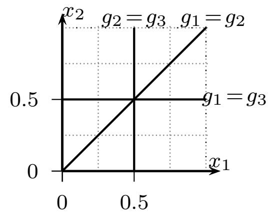
A) $\{\mathbf{x} : x_1 \geq 0 \wedge x_2 < -x_1 + 2\} \cup \{\mathbf{x} : x_1 < 0 \wedge x_2 < x_1 + 2\}$.
B) $\{\mathbf{x} : x_2 \geq 0 \wedge x_2 < -x_1 + 2\} \cup \{\mathbf{x} : x_2 < 0 \wedge x_2 > x_1 - 2\}$.
C) $\{\mathbf{x} : x_1 \geq 0 \wedge x_2 < -x_1 + 1\} \cup \{\mathbf{x} : x_1 < 0 \wedge x_2 < x_1 + 1\}$.
D) $\{\mathbf{x} : x_2 \geq 0 \wedge x_2 < -x_1 + 1\} \cup \{\mathbf{x} : x_2 < 0 \wedge x_2 > x_1 - 1\}$.

39 ☐ En la figura se representan 4 muestras de aprendizaje de sendas clases: $\mathbf{x}_1 = (1,1)^t$ de la clase $c_1 = 1$, $\mathbf{x}_2 = (-1,1)^t$ de $c_2 = 2$, $\mathbf{x}_3 = (-1,-1)^t$ de $c_3 = 3$, y $\mathbf{x}_4 = (1,-1)^t$ de $c_4 = 4$. Supóngase que se ejecuta el algoritmo Perceptrón a partir de las mismas, con factor de aprendizaje $\alpha = 1$, margen $b = 0.1$ y vectores de pesos iniciales nulos (en notación homogénea). Durante la primera iteración del algoritmo y tras procesar las 3 primeras muestras, se obtienen los vectores de pesos $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12})^t = (0,2,0)^t$, $\mathbf{w}_2 = (-1,-1,1)^t$, $\mathbf{w}_3 = (-1,-1,-3)^t$ y $\mathbf{w}_4 = (-3,1,-1)^t$. Completa la primera iteración del algoritmo e indica, a partir de los vectores de pesos resultantes, cuántas muestras de aprendizaje se clasifican *correctamente:*



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

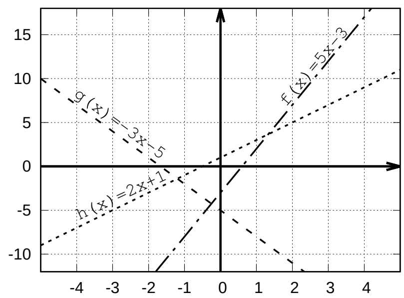
40 ☐ En la figura de la derecha se representan las fronteras de decisión de un clasificador en 3 clases. ¿Cuales de los siguientes vectores de pesos definen dichas fronteras?



A) $\mathbf{w}_1 = (0,0,1)^t$ $\mathbf{w}_2 = (0,1,0)^t$ y $\mathbf{w}_3 = (1,0,0)^t$
B) $\mathbf{w}_1 = (0,0,1)^t$ $\mathbf{w}_2 = (0,1,0)^t$ y $\mathbf{w}_3 = (0.5,0,0)^t$
C) $\mathbf{w}_1 = (0.5,0,0)^t$ $\mathbf{w}_2 = (0,1,0)^t$ y $\mathbf{w}_3 = (0,0,1)^t$
D) $\mathbf{w}_1 = (0,0,1)^t$ $\mathbf{w}_2 = (1,0,0)^t$ y $\mathbf{w}_3 = (0,1,0)^t$

41 ☐ Sea un clasificador lineal para dos clases, $\circ$ y $\bullet$, de vectores de pesos $\mathbf{a}_\circ = (2,-5,4)^t$ y $\mathbf{a}_\bullet = (5,1,1)^t$, respectivamente. ¿Cuál de las siguientes afirmaciones es correcta?
A) Los vectores de pesos $\mathbf{a}_\circ = (3,4,1)^t$ y $\mathbf{a}_\bullet = (2,2,2)^t$ definen la misma frontera de decisión que los del enunciado.
B) Los vectores de pesos $\mathbf{a}_\circ = (-2,5,-4)^t$ y $\mathbf{a}_\bullet = (-5,-1,-1)^t$ definen un clasificador equivalente al del enunciado.
C) El punto $\mathbf{x}' = (1,2)^t$ pertenece a la clase $\circ$.
D) El punto $\mathbf{x}' = (-2,0)^t$ se encuentra en la frontera de decisión.

42 ☐ En la figura de la derecha se muestran las funciones discriminantes lineales resultantes de entrenar un clasificador con el algoritmo Perceptrón con un conjunto de puntos de $\mathbb{R}$. Las funciones obtenidas son: $g(x) = -3x - 5$, $h(x) = 2x + 1$ y $f(x) = 5x - 3$. Indica cuáles son las fronteras de decisión correctas entre $g(x)$ y $h(x)$, y entre $h(x)$ y $f(x)$:



A) $x = -5/3$ y $x = 3/5$.
B) $x = -1/2$ y $x = 3/5$.
C) $x = -6/5$ y $x = 4/3$.
D) $x = -5/3$ y $x = 4/3$.

43 ☐ Indica cuál de las siguientes afirmaciones referentes al algoritmo Perceptrón (al que llamaremos P) es *cierta* cuando se aplica al aprendizaje con una muestra de vectores etiquetados S:
A) Si la muestra de aprendizaje es linealmente separable, P termina tras un número finito de iteraciones y los pesos resultantes permiten clasificar S sin errores.
B) El número de vectores de S bien clasificados con los pesos obtenidos en cada iteración de P es mayor que el número vectores bien clasificados en la iteración anterior.
C) P siempre converge en un número finito de iteraciones, aunque es posible que los pesos finalmente obtenidos no clasifiquen correctamente a todos los vectores de S.
D) Cuanto más grande es S, mayor es el número de iteraciones que necesita P para converger.

44 ☐ Sea un problema de clasificación en cuatro clases de objetos representados en $\mathbb{R}^3$. Se tiene un clasificador cuyas funciones discriminantes son lineales con vectores de pesos (en notación homogénea):

$$\mathbf{a}_1 = (-2,1,2,0)^t \qquad \mathbf{a}_2 = (0,2,2,0)^t \qquad \mathbf{a}_3 = (1,1,1,0)^t \qquad \mathbf{a}_4 = (3,0,0,2)^t$$

Indica a qué clase se asignarà el objeto $\mathbf{x} = (1,2,2)^t$ (*no* en notación homógenea).
A) 1.
B) 2.
C) 3.
D) 4.

45 ☐ Supóngase que se está aplicando el algoritmo Perceptrón con $b = 1.5$ y que los vectores de pesos actuales de las clases son los dados en la cuestión anterior. Asimismo, supóngase que el objeto $\mathbf{x} = (1, 2, 2)^t$ dado en la cuestión anterior es la siguiente muestra de entrenamiento a procesar, la cual suponemos perteneciente a la clase 3. Entonces:

A) Se modificarán los vectores de pesos $\mathbf{a}_2$, $\mathbf{a}_3$ y $\mathbf{a}_4$.
B) Se modificará sólo el vector de pesos $\mathbf{a}_3$.
C) No se modificará ningún vector de pesos.
D) Se modificarán todos los vectores de pesos.

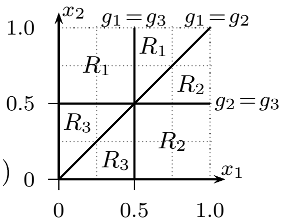
46 ☐ La notación homogénea se usa para poder escribir funciones discriminantes lineales $g(\cdot)$ en forma vectorial compacta. Sean: E un espacio de representación de dimensión 3; $\mathbf{y} \in E$ un punto de E; a_0, a_1, a_2 y a_3 cuatro coeficientes reales; $\mathbf{w} \stackrel{\text{def}}{=} (a_1, a_2, a_3)^t$ un vector real de dimensión 3; y $\mathbf{a} \stackrel{\text{def}}{=} (a_0, a_1, a_2, a_3)^t$ un vector real de dimensión 4. Indica cuál de las siguientes expresiones hace un uso *incorrecto* de la notación homogénea:

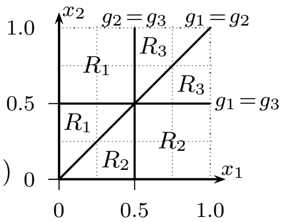
A) $g(\mathbf{y}) = \mathbf{a}^t\, \mathbf{y}$
B) $g(\mathbf{y}) = \mathbf{w}^t\, \mathbf{y} + a_0$
C) $g(\mathbf{y}) = \mathbf{a}^t\, \mathbf{x}$, donde $\mathbf{x} \stackrel{\text{def}}{=} (1, y_1, y_2, y_3)^t$
D) $g(\mathbf{y}) = a_0 + (a_1, a_2, a_3)^t\, \mathbf{y}$

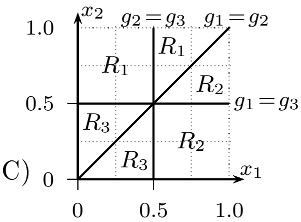
47 ☐ Sea $S = \{(\mathbf{y}_1, c_1), \ldots, (\mathbf{y}_N, c_N)\}, 1 \leq c_j \leq C, 1 \leq j \leq N$, un conjunto *linealmente separable* de muestras representativas de aprendizaje en notación homogénea. S se usa como entrada al algoritmo *Perceptrón*, inicializado con $\mathbf{a}'_j = \mathbf{0}$, $1 \leq j \leq C$. Tras un número suficientemente grande de iteraciones con un factor de aprendizaje $\alpha = 1$ y margen $b = 10$, el algoritmo termina y obtiene C vectores de pesos, en notación homogénea, $\mathbf{a}_j,\ 1 \leq j \leq C$. Indica cuál de las siguienes afirmaciones es *correcta*.

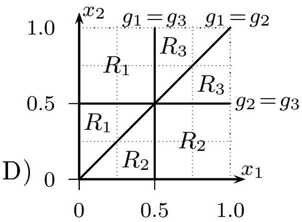
A) Se cumplen las $N \cdot C$ inecuaciones:
$\mathbf{a}_i^t \mathbf{y}_i > \mathbf{a}_j^t \mathbf{y}_j\ \ 1 \leq i \leq N,\ 1 \leq j \leq C,\ i \neq j$
B) Las N muestras de S se clasifican correctamente; es decir, se cumplen las N inecuaciones:
$\mathbf{a}_{c_i}^t \mathbf{y}_i > \mathbf{a}_j^t \mathbf{y}_i,\ \ 1 \leq i \leq N,\ j \neq c_i$
C) Las N muestras de S se clasifican correctamente, es decir, se cumplen las $N \cdot (C-1)$ inecuaciones:
$\mathbf{a}_{c_i}^t \mathbf{y}_i > \mathbf{a}_j^t \mathbf{y}_i,\ \ 1 \leq i \leq N,\ 1 \leq j \leq C,\ j \neq c_i$
D) Aunque S es separable, como $b \gg \alpha > 0$, no se puede afirmar que todas las muestras de S se clasifiquen correctamente con los vectores de pesos $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_C$.

48 ☐ Sea un clasificador en tres clases basado en las funciones discriminantes lineales bidimensionales de vectores de pesos: $\mathbf{w}_1 = (0, 0, 1)^t$, $\mathbf{w}_2 = (0, 1, 0)^t$ y $\mathbf{w}_3 = (0.5, 0, 0)^t$. Indica cuál de las figuras dadas a continuación es coherente con las fronteras y regiones de decisión que define dicho clasificador.

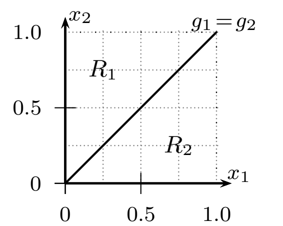
A) 

B) 

C) 

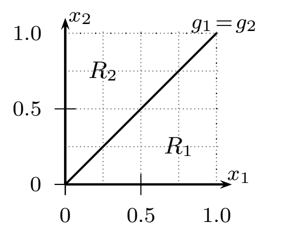
D) 

49 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos *no* define un clasificador equivalente al dado?



A) $\mathbf{w}_1 = (0, 0, 1)^t$ y $\mathbf{w}_2 = (0, 1, 0)^t$.
B) $\mathbf{w}_1 = (0, 1, 0)^t$ y $\mathbf{w}_2 = (0, 0, 1)^t$.
C) $\mathbf{w}_1 = (0, -1, 0)^t$ y $\mathbf{w}_2 = (0, 0, -1)^t$.
D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

50 ☐ Durante la aplicación del algoritmo Perceptrón ($\alpha = 1.0$ y $b = 0$) en un problema de clasificación en dos clases, se han obtenido los vectores de pesos $\mathbf{w}_1 = (-1, 1, 0)^t$ y $\mathbf{w}_2 = (1, 0, 1)^t$. Supón que el siguiente paso en la aplicación de Perceptrón consiste en procesar una cierta muestra de entrenamiento $\mathbf{x}$ de clase c. Indica cuál de las siguientes opciones daría como resultado un conjunto de pesos que define la frontera y regiones de decisión de la figura de la derecha.



A) $\mathbf{x} = (-1, 1)^t$ y $c = 2$.
B) $\mathbf{x} = (0, 0)^t$ y $c = 2$.
C) $\mathbf{x} = (-1, 1)^t$ y $c = 1$
D) $\mathbf{x} = (0, 0)^t$ y $c = 1$.

51 ☐ Sea un problema de clasificación en tres clases para objetos del tipo $\mathbf{x} = (x_1, x_2)^t \in \{0,1\}^2$, con las distribuciones de probabilidad de la derecha. ¿Cuál es el error de Bayes, ε^*, en este problema?

$\mathbf{x}$		$P(c \mid \mathbf{x})$			
x_1	x_2	$c{=}1$	$c{=}2$	$c{=}3$	$P(\mathbf{x})$
0	0	0.6	0.2	0.2	0.2
0	1	0.1	0.1	0.8	0.3
1	0	0.3	0.5	0.2	0.2
1	1	1/3	1/3	1/3	0.3

A) $\varepsilon^* < 0.2$.
B) $0.2 \leq \varepsilon^* < 0.4$.
C) $0.4 \leq \varepsilon^* < 0.7$.
D) $0.7 \leq \varepsilon^*$.

52 ☐ Se tiene un problema de clasificación para el cual se ha aprendido un clasificador. Asimismo, se tiene un conjunto de $M = 100$ muestras de test con el cual se ha estimado:

- La probabilidad de error del clasificador aprendido, $\hat{p} = 0.10 = 10\,\%$.
- Un intervalo de confianza al 95 % para dicha probabilidad de error, $\hat{I} = [0.04, 0.16] = [4\,\%, 16\,\%]$.

Se considera que la probabilidad de error estimada es razonable y que la misma no variará significativamente aunque usemos muchas más muestras de test. Ahora bien, el intervalo de confianza (al 95 %) estimado, $\hat{I} = 10\,\% \pm 6\,\%$, nos parece un poco amplio y nos preguntamos si es posible reducir su amplitud mediante el uso de más de $M = 100$ muestras de test. Además, si ello fuera posible, nos preguntamos si sería posible reducir dicha amplitud a la mitad o menos; esto es, tal que $\hat{I} = 10\,\% \pm \hat{R}$ con $\hat{R} \leq 3\,\%$. En relación con estas cuestiones, indica cuál de las siguientes afirmaciones es correcta.

A) En general, no es posible reducir la amplitud de $\hat{I}$ pues $\hat{I}$ no depende significativamente de M.
B) No es posible reducir la amplitud de $\hat{I}$ ya que hemos considerado que $\hat{p}$ no variará significativamente y, siendo así, la amplitud de $\hat{I}$ tampoco puede variar significativamente.
C) Sí es posible reducir la amplitud de $\hat{I}$, a la mitad o menos, si doblamos M al menos ($M \geq 200$).
D) Sí es posible reducir la amplitud de $\hat{I}$, a la mitad o menos, si empleamos al menos cuatro veces más muestras de test aproximadamente ($M \geq 400$).

53 ☐ La expresión $\hat{c} = \arg\max_{1\leq c\leq C} P(c \mid \mathbf{y})$, donde $\mathbf{y} \in \mathbb{R}^d$ es un dato a clasificar, corresponde a un clasificador de mínimo riesgo de error o de Bayes en C clases. Con algunas asunciones, este clasificador coincide con un clasificador basado en C *Funciones Discriminantes*, definido por $\hat{c} = \arg\max_{1\leq c\leq C} g_c(\mathbf{y})$. Indica cuál de las siguientes asunciones *no* sería generalmente correcta:

A) $g_c(\mathbf{y}) = P(c \mid \mathbf{y})$.
B) $g_c(\mathbf{y}) = \log P(c \mid \mathbf{y})$.
C) $g_c(\mathbf{y}) = 0.5 \cdot P(c \mid \mathbf{y}) + 0.5$.
D) $g_c(\mathbf{y}) = \sum_{j=1}^{d} a_j P(c \mid y_j) + a_0$, donde $a_j, 0 \leq j \leq d$, son coeficientes reales no nulos cualesquiera.

54 ☐ Sea S un conjunto de N pares de entrenamiento y C el número de clases. Considera una iteración cualquiera, que no sea la última, del algoritmo Perceptrón aplicado a S. En esa iteración se modifican k vectores de pesos. ¿Cuál de las siguientes afirmaciones es *incorrecta*?

A) $2 \leq k \leq C \cdot N$.
B) $2 \leq k \leq (C-1) \cdot N$.
C) $2 \leq k \leq C' \cdot N$, donde C' está acotada según $1 \leq C' \leq C$.
D) $2 \leq k \leq \sum_{n=1}^{N} k_n$, donde $k_n, 1 \leq k_n \leq C$, es el número de vectores modificados para el dato n-ésimo.

55 ☐ La probabilidad de error de un clasificador se estima que es del 14 %. Determina cuál es el número mínimo de muestras de test necesario, M, para conseguir que el intervalo de confianza al 95 % de dicho error no supere el $\pm 1\,\%$; esto es, $I = [13\,\%, 15\,\%]$:

A) $M < 2000$.
B) $2000 \leq M < 3500$.
C) $3500 \leq M < 5000$.
D) $M \geq 5000$.

56 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0,0,0)^t$, $\mathbf{w}_2 = (0,0,0)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los vectores de pesos $\mathbf{w}_1 = (-1,-5,-5)^t$, $\mathbf{w}_2 = (1,5,5)^t$, ¿cuál de las siguientes es esa última muestra?

A) $((5,5)^t, 2)$

B) $((3,5)^t, 2)$
C) $((5,1)^t, 2)$
D) $((3,1)^t, 2)$

57 ☐ Dado el clasificador en 2 clases definido por sus vectores de pesos $\mathbf{w}_1 = (-1,2,2)^t$, $\mathbf{w}_2 = (-2,-3,3)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1 = (-1,6,6)^t$, $\mathbf{w}_2 = (-4,-9,9)^t$
B) $\mathbf{w}_1 = (2,-4,-4)^t$, $\mathbf{w}_2 = (4,6,-6)^t$
C) $\mathbf{w}_1 = (1,2,2)^t$, $\mathbf{w}_2 = (0,-3,3)^t$
D) $\mathbf{w}_1 = (-3,6,6)^t$, $\mathbf{w}_2 = (-6,-9,9)^t$

58 ☐ Dado el clasificador en 2 clases definido por sus vectores de pesos $\mathbf{w}_1 = (0,-3,1,1)^t$, $\mathbf{w}_2 = (0,1,0,2)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1 = (1,-3,1,1)^t$, $\mathbf{w}_2 = (1,1,0,2)^t$
B) $\mathbf{w}_1 = (0,-9,3,3)^t$, $\mathbf{w}_2 = (0,3,0,6)^t$
C) $\mathbf{w}_1 = (0,6,-2,-2)^t$, $\mathbf{w}_2 = (0,-2,0,-4)^t$
D) $\mathbf{w}_1 = (1,-9,3,3)^t$, $\mathbf{w}_2 = (1,3,0,6)^t$

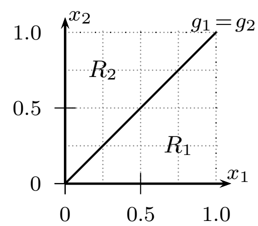
59 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0,-1,-1)^t$, $\mathbf{w}_2 = (0,1,1)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los vectores de pesos $\mathbf{w}_1 = (1,0,3)^t$, $\mathbf{w}_2 = (-1,0,-3)^t$, ¿cuál de las siguientes es esa última muestra?

A) $((1,4)^t, 1)$
B) $((5,4)^t, 1)$
C) $((3,2)^t, 1)$
D) $((4,5)^t, 2)$

60 ☐ La probabilidad de error de un clasificador se estima que es del 3 %. Determina cuál es el número mínimo de muestras de test necesario, M, para conseguir que el intervalo de confianza al 95 % de dicho error no supere el ± 1 %; esto es, $I = [2\,\%, 4\,\%]$:

A) $M < 2000$.
B) $2000 \leq M < 3500$.
C) $3500 \leq M < 5000$.
D) $M \geq 5000$.

61 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos (en notación homogénea) define un clasificador equivalente al dado?

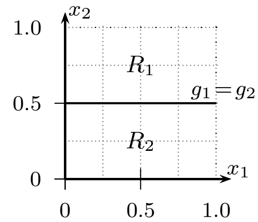


A) $\mathbf{w}_1 = (0,-2,0)^t$ y $\mathbf{w}_2 = (0,0,-2)^t$.
B) $\mathbf{w}_1 = (0,2,0)^t$ y $\mathbf{w}_2 = (0,0,2)^t$.
C) $\mathbf{w}_1 = (0,0,2)^t$ y $\mathbf{w}_2 = (0,2,0)^t$.
D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

62 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 4 muestras bidimensionales de aprendizaje para un problema de 4 clases, $c = 1,2,3,4$. En un momento dado de la ejecución del algoritmo se han obtenido los vectores de pesos $\mathbf{w}_1 = (-2,-2,-6)^t$, $\mathbf{w}_2 = (-2,-2,-6)^t$, $\mathbf{w}_3 = (-2,-4,-4)^t$, $\mathbf{w}_4 = (-2,-4,-4)^t$. Suponiendo que a continuación se va a procesar la muestra $(\mathbf{x}, c) = ((4,5)^t, 2)$, ¿cuántos vectores de pesos se modificarán?

A) 0
B) 2
C) 3
D) 4

63 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos (en notación homogénea) define un clasificador equivalente al dado?



A) $\mathbf{w}_1 = (0, 0, 2)^t$ y $\mathbf{w}_2 = (1, 0, 0)^t$.
B) $\mathbf{w}_1 = (0, 0, -2)^t$ y $\mathbf{w}_2 = (-1, 0, 0)^t$.
C) $\mathbf{w}_1 = (1, 0, 0)^t$ y $\mathbf{w}_2 = (0, 0, 2)^t$.
D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

64 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 4 muestras bidimensionales de aprendizaje para un problema de 4 clases, $c = 1, 2, 3, 4$. En un momento dado de la ejecución del algoritmo se han obtenido los vectores de pesos $\mathbf{w}_1 = (-2, -8, -12)^t$, $\mathbf{w}_2 = (-2, -6, -4)^t$, $\mathbf{w}_3 = (-2, -6, -4)^t$, $\mathbf{w}_4 = (-2, -8, -8)^t$. Suponiendo que a continuación se va a procesar la muestra $(\mathbf{x}, c) = ((5, 3)^t, 2)$, ¿cuántos vectores de pesos se modificarán?

A) 0
B) 2
C) 3
D) 4

65 ☐ Dado el clasificador en 2 clases definido por sus vectores de pesos $\mathbf{w}_1 = (1, 1, 2)^t$, $\mathbf{w}_2 = (2, 0, 3)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1 = (-1, -1, -2)^t$, $\mathbf{w}_2 = (-2, 0, -3)^t$
B) $\mathbf{w}_1 = (2, 1, 2)^t$, $\mathbf{w}_2 = (3, 0, 3)^t$
C) $\mathbf{w}_1 = (3, 2, 4)^t$, $\mathbf{w}_2 = (5, 0, 6)^t$
D) $\mathbf{w}_1 = (2, 2, 4)^t$, $\mathbf{w}_2 = (4, 0, 6)^t$

66 ☐ Dado el clasificador en 2 clases definido por sus vectores de pesos $\mathbf{w}_1 = (-2, 3, 3)^t$, $\mathbf{w}_2 = (0, 2, -2)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1 = (1, 3, 3)^t$, $\mathbf{w}_2 = (3, 2, -2)^t$
B) $\mathbf{w}_1 = (-4, 6, 6)^t$, $\mathbf{w}_2 = (0, 4, -4)^t$
C) $\mathbf{w}_1 = (-1, 6, 6)^t$, $\mathbf{w}_2 = (3, 4, -4)^t$
D) $\mathbf{w}_1 = (2, -3, -3)^t$, $\mathbf{w}_2 = (0, -2, 2)^t$

67 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0, 0, -2)^t$, $\mathbf{w}_2 = (0, 0, 2)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los vectores de pesos $\mathbf{w}_1 = (1, 1, -1)^t$, $\mathbf{w}_2 = (-1, -1, 1)^t$, ¿cuál de las siguientes es esa última muestra?

A) $((2, 3)^t, 1)$
B) $((1, 1)^t, 1)$
C) $((2, 1)^t, 2)$
D) $((2, 5)^t, 2)$

68 ☐ Dado el clasificador en 2 clases definido por sus vectores de pesos $\mathbf{w}_1 = (-1, 3, 1, -3)^t$, $\mathbf{w}_2 = (-3, -2, 2, 2)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1 = (0, 3, 1, -3)^t$, $\mathbf{w}_2 = (-2, -2, 2, 2)^t$
B) $\mathbf{w}_1 = (-2, 9, 3, -9)^t$, $\mathbf{w}_2 = (-8, -6, 6, 6)^t$
C) $\mathbf{w}_1 = (-3, 9, 3, -9)^t$, $\mathbf{w}_2 = (-9, -6, 6, 6)^t$
D) $\mathbf{w}_1 = (2, -6, -2, 6)^t$, $\mathbf{w}_2 = (6, 4, -4, -4)^t$

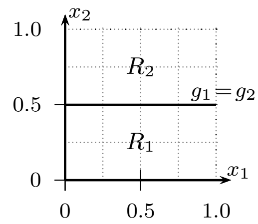
69 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$, a un conjunto de 3 muestras bidimensionales de aprendizaje para un problema de 2 clases. Tras procesar las primeras 2 muestras se han obtenido los vectores de pesos $\mathbf{w}_1 = (0, 2, 1)^t$, $\mathbf{w}_2 = (0, -2, -1)^t$. A continuación, se procesa la última muestra $(\mathbf{x}_3, c_3)$ y se obtienen los vectores de pesos $\mathbf{w}_1 = (-1, 1, -3)^t$, $\mathbf{w}_2 = (1, -1, 3)^t$, ¿cuál de las siguientes es esa última muestra?

A) $((2,1)^t,1)$
B) $((3,1)^t,1)$
C) $((1,4)^t,2)$
D) $((2,4)^t,1)$

70 ☐ Dado el clasificador en 3 clases definido por sus vectores de pesos $\mathbf{w}_1=(2,1,1)^t$, $\mathbf{w}_2=(1,-3,-3)^t$, $\mathbf{w}_3=(2,0,-1)^t$ en notación homogénea, ¿cuál de los siguientes conjuntos de vectores **no** define un clasificador equivalente al dado?

A) $\mathbf{w}_1=(-2,-1,-1)^t$, $\mathbf{w}_2=(-1,3,3)^t$, $\mathbf{w}_3=(-2,0,1)^t$
B) $\mathbf{w}_1=(4,2,2)^t$, $\mathbf{w}_2=(2,-6,-6)^t$, $\mathbf{w}_3=(4,0,-2)^t$
C) $\mathbf{w}_1=(4,1,1)^t$, $\mathbf{w}_2=(3,-3,-3)^t$, $\mathbf{w}_3=(4,0,-1)^t$
D) $\mathbf{w}_1=(6,2,2)^t$, $\mathbf{w}_2=(4,-6,-6)^t$, $\mathbf{w}_3=(6,0,-2)^t$

71 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos (en notación homogénea) define un clasificador equivalente al dado?



A) $\mathbf{w}_1=(-1,0,0)^t$ y $\mathbf{w}_2=(0,0,-2)^t$.
B) $\mathbf{w}_1=(1,0,0)^t$ y $\mathbf{w}_2=(0,0,2)^t$.
C) $\mathbf{w}_1=(0,0,2)^t$ y $\mathbf{w}_2=(1,0,0)^t$.
D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

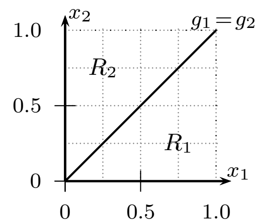
72 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha=1$ y margen $b=0.1$, a un conjunto de 4 muestras bidimensionales de aprendizaje para un problema de 4 clases, $c=1,2,3,4$. En un momento dado de la ejecución del algoritmo se han obtenido los vectores de pesos $\mathbf{w}_1=(-2,-3,-9)^t$, $\mathbf{w}_2=(-2,-5,-5)^t$, $\mathbf{w}_3=(-2,-7,-11)^t$, $\mathbf{w}_4=(-2,-3,-5)^t$. Suponiendo que a continuación se va a procesar la muestra $(\mathbf{x},c)=((3,4)^t,3)$, ¿cuántos vectores de pesos se modificarán?

A) 0
B) 2
C) 3
D) 4

73 ☐ La probabilidad de error de un clasificador se estima que es del 7 %. Determina cuál es el número mínimo de muestras de test necesario, M, para conseguir que el intervalo de confianza al 95 % de dicho error no supere el $\pm 1\,\%$; esto es, $I=[6\,\%,8\,\%]$:

A) $M<1000$.
B) $1000\le M<2000$.
C) $2000\le M<3000$.
D) $M\ge 3000$.

74 ☐ Dado el clasificador en dos clases definido por su frontera y regiones de decisión de la figura de la derecha, ¿cuál de los siguientes vectores de pesos (en notación homogénea) define un clasificador equivalente al dado?



A) $\mathbf{w}_1=(0,1,0)^t$ y $\mathbf{w}_2=(0,0,1)^t$.
B) $\mathbf{w}_1=(0,-1,0)^t$ y $\mathbf{w}_2=(0,0,-1)^t$.
C) $\mathbf{w}_1=(0,0,1)^t$ y $\mathbf{w}_2=(0,1,0)^t$.
D) Todos los vectores de pesos anteriores definen clasificadores equivalentes.

75 ☐ Supóngase que estamos aplicando el algoritmo Perceptrón, con factor de aprendizaje $\alpha=1$ y margen $b=0.1$, a un conjunto de 4 muestras bidimensionales de aprendizaje para un problema de 4 clases, $c=1,2,3,4$. En un momento dado de la ejecución del algoritmo se han obtenido los vectores de pesos $\mathbf{w}_1=(-2,-8,-5)^t$, $\mathbf{w}_2=(-2,-8,-9)^t$, $\mathbf{w}_3=(-2,0,-3)^t$, $\mathbf{w}_4=(-2,-4,-9)^t$. Suponiendo que a continuación se va a procesar la muestra $(\mathbf{x},c)=((5,4)^t,1)$, ¿cuántos vectores de pesos se modificarán?

A) 0
B) 2

C) 3
D) 4

76 ☐ La probabilidad de error de un clasificador se estima que es del 5 %. Determina cuál es el número mínimo de muestras de test necesario, M, para conseguir que el intervalo de confianza al 95 % de dicho error no supere el $\pm 1\,\%$; esto es, $I = [4\,\%, 6\,\%]$:

A) $M < 1000$.
B) $1000 \leq M < 2000$.
C) $2000 \leq M < 3000$.
D) $M \geq 3000$.

2. Problemas

1. Sea un problema de clasificación en tres clases, $c = \{A, B, C\}$, para objetos representados mediante vectores de características tridimensionales. Se tiene un clasificador lineal basado en funciones discriminantes lineales de la forma

$$g_c(\mathbf{x}) = \mathbf{w}_c \cdot \mathbf{x} \qquad \text{para toda clase } c$$

donde $\mathbf{w}_c$ y $\mathbf{x}$ se hallan en notación compacta; es decir: $\mathbf{w} = (w_0, w_1, w_2, w_3)^t \in \mathbb{R}^4$ y $\mathbf{x} = (x_0, x_1, x_2, x_3)^t \in \mathbb{R}^4$, con $x_0 = 1$. Teniendo en cuenta que:

$$\mathbf{w}_A = (1, 1, 1, 1)^t \qquad \mathbf{w}_B = (-1, 0, -1, -2)^t \qquad \text{y} \qquad \mathbf{w}_C = (-2, 2, -1, 0)^t$$

Se pide:

a) Clasifica el punto $\mathbf{x}' = (2, 1, 2)^t$.
b) Sabemos que el punto $\mathbf{x}' = (-1, 0, -1)^t$ pertenece a la clase A. ¿Qué valores tendrían $\mathbf{w}_A$,$\mathbf{w}_B$ y $\mathbf{w}_C$ tras aplicar el algoritmo Perceptrón para dicho punto, usando una constante de aprendizaje $\alpha = 0.1$?
c) Dado el punto $\mathbf{x}' = (1, -1, 2)^t$ que pertenece a la clase A, obtén valores posibles de las discriminantes que lo clasifiquen correctamente.

2. Sea un problema de clasificación en 3 clases, $c = \{1, 2, 3\}$, para objetos representados mediante vectores de características bidimensionales. Se tienen 3 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, $\mathbf{x}_2 = (0, 1)^t$ de $c_2 = 2$, y $\mathbf{x}_3 = (2, 2)^t$ de $c_3 = 3$. Encuentra un clasificador lineal de mínimo error mediante el algoritmo Perceptrón, con vectores de pesos iniciales de las clases nulos, factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$. Presenta una traza de ejecución que incluya las sucesivas actualizaciones de los vectores de pesos de las clases.

3. Sea un problema de clasificación en 2 clases, $c = \{1, 2\}$, para objetos representados mediante vectores de características bidimensionales. Se tienen 2 muestras de entrenamiento: $\mathbf{x}_1 = (0, 0)^t$ de la clase $c_1 = 1$, y $\mathbf{x}_2 = (1, 1)^t$ de $c_2 = 2$. Encuentra un clasificador lineal de mínimo error mediante el algoritmo Perceptrón, con vectores de pesos iniciales de las clases nulos, factor de aprendizaje $\alpha = 1$ y margen $b = 0.1$. Presenta una traza de ejecución que incluya las sucesivas actualizaciones de los vectores de pesos de las clases.

4. (Examen de SIN del 18 de enero de 2013; tiempo estimado: 30 minutos)
En un problema de clasificación en 2 clases, A, B, para objetos representados mediante vectores de características bidimensionales, se tienen dos muestras de entrenamiento:

$$\mathbf{y}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \in A, \quad \mathbf{y}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \in B,$$

a) Inicializando a 0 todas las componentes de los vectores de pesos iniciales, mostrar una traza de ejecución del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1.0$ y margen $b = 0.1$. La traza debe incluir las sucesivas actualizaciones de los vectores de pesos de las clases e indicar los vectores de pesos obtenidos como solución final.
b) Obtener la ecuación de la frontera de separación entre clases correspondiente a la solución obtenida. Representar gráficamente esta frontera, junto con los datos de entrenamiento. ¿Es satisfactoria esta solución?

5. En la tabla siguiente se proporciona un conjunto de 4 muestras bidimensionales de aprendizaje de 3 clases, $c = 1, 2, 3$.

n	x_{n1}	x_{n2}	c_n
1	5	2	1
2	1	2	3
3	1	1	2
4	4	1	1

Se pide:

a) (1.5 puntos) Realiza una traza de ejecución de una iteración del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$, margen $b = 0.1$ y los siguientes pesos iniciales de cada clase por columnas:

d	w_{d1}	w_{d2}	w_{d3}
0	-3	0	-1
1	1	-4	-5
2	-4	-2	0

b) (0.5 puntos) Clasifica la muestra de test $\mathbf{x} = (5, 5)^t$ mediante un clasificador lineal con los vectores de pesos obtenidos en el apartado anterior.

6. En la tabla siguiente se proporciona un conjunto de 4 muestras bidimensionales de aprendizaje de 3 clases, $c = 1, 2, 3$.

n	x_{n1}	x_{n2}	c_n
1	3	5	3
2	5	1	1
3	2	2	1
4	1	2	2

Se pide:

a) (1.5 puntos) Realiza una traza de ejecución de una iteración del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$, margen $b = 0.1$ y los siguientes pesos iniciales de cada clase por columnas:

d	w_{d1}	w_{d2}	w_{d3}
0	-1	0	-3
1	4	-6	-8
2	-10	-2	2

b) (0.5 puntos) Clasifica la muestra de test $\mathbf{x} = (5, 5)^t$ mediante un clasificador lineal con los vectores de pesos obtenidos en el apartado anterior.

7. Se tiene un problema de clasificación en dos clases, 0 y 1, para objetos representados en $\{0, 1\}^2$, esto es, vectores de bits de la forma $\mathbf{x} = (x_1, x_2)^t$ con $x_1, x_2 \in \{0, 1\}$. Asimismo, disponemos de cuatro muestras de entrenamiento:

$\mathbf{x}_n$	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$
x_{n1}	0	0	1	1
x_{n2}	0	1	0	1
c_n	0	1	1	0

Se pide:

1 (0.75 puntos) Aplica una iteración del algoritmo Perceptrón con pesos iniciales nulos, constante de aprendizaje $\alpha = 1$ y margen $b = 0.1$. ¿Qué pesos se obtienen al finalizar la iteración aplicada?

2 (0.50 puntos) A partir de la inicialización dada en el apartado anterior, ¿convergerá el algoritmo Perceptrón a una solución sin datos de entrenamiento mal clasificados?
Indica si sí o no y después comenta brevemente la respuesta.

3 (0.25 puntos) A partir de alguna inicialización con pesos no nulos, $\alpha > 0$ y $b = 0.1$, ¿convergerá el algoritmo Perceptrón a una solución sin datos de entrenamiento mal clasificados?
Indica si sí o no y después comenta brevemente la respuesta.

8. En la tabla de la izquierda se proporciona un conjunto de 3 muestras bidimensionales de aprendizaje de 3 clases, mientras que en la tabla de la derecha se proporciona un conjunto de pesos iniciales para cada clase.

n	x_{n1}	x_{n2}	c_n
1	-2	-2	1
2	0	0	2
3	2	2	3

	$\mathbf{w}_1$	$\mathbf{w}_2$	$\mathbf{w}_3$
w_{c0}	0	-1	-1
w_{c1}	-2	0	4
w_{c2}	-2	0	4

Se pide:

a) (1.5 puntos) Realiza una traza de ejecución de una iteración del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$, margen $\gamma = 0.1$ utilizando los pesos iniciales proporcionados.

b) (0.5 puntos) Representa gráficamente las regiones de decisión del clasificador resultante, así como las fronteras de decisión necesarias para su representación.

9. En la tabla de la izquierda se proporciona un conjunto de 3 muestras bidimensionales de aprendizaje de 3 clases, mientras que en la tabla de la derecha se proporciona un conjunto de pesos iniciales para cada clase.

n	x_{n1}	x_{n2}	c_n
1	1	0	1
2	0	0	2
3	-1	0	3

	$\mathbf{w}_1$	$\mathbf{w}_2$	$\mathbf{w}_3$
w_{c0}	-2	-1	-1
w_{c1}	2	1	-3
w_{c2}	0	0	0

Se pide:

a) (1.5 puntos) Realiza una traza de ejecución de una iteración del algoritmo Perceptrón, con factor de aprendizaje $\alpha = 1$, margen $\gamma = 0.1$ utilizando los pesos iniciales proporcionados.

b) (0.5 puntos) Representa gráficamente las regiones de decisión del clasificador resultante, así como las fronteras de decisión necesarias para su representación.