Examen del bloc 2 de SIN (tipus A)

ETSINF, Universitat Politècnica de València, 14 de gener de 2020

Cognoms: ______________________ **Nom:** ______________

Grup: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3E** ☐ **3F** ☐ **3G** ☐ **4IA**

Test (1,75 punts)

Marca cada requadre amb una única opció. Puntuació: $\max(0, (\text{encerts} - \text{errors}\,/\,3) \cdot 1,75\,/\,9)$.

1 ☐ Siga $\mathbf{x}$ un objecte a classificar en una classe de C possibles. Indica quin dels següents classificadors *no* és d'error mínim (o tria l'última opció si els tres són d'error mínim):

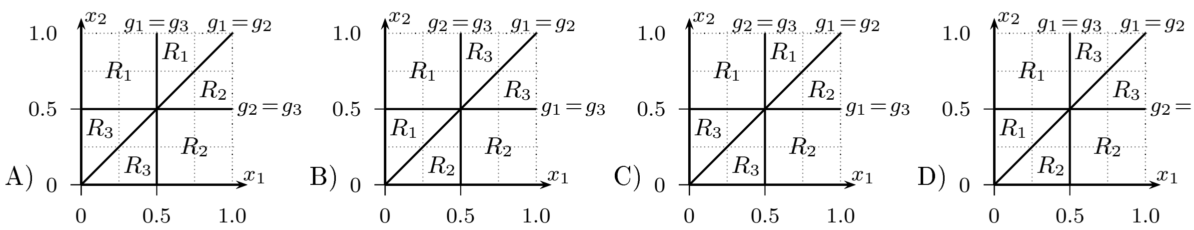
A) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\max}\; p(c \mid \mathbf{x})^2$.

B) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\max}\; \log p(\mathbf{x}, c)$.

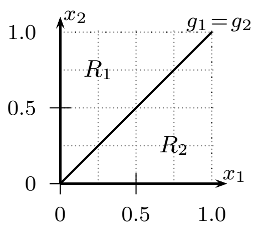
C) $c(\mathbf{x}) = \underset{c=1,\dots,C}{\arg\max}\; \sqrt{p(\mathbf{x}, c)}\,/\,p(\mathbf{x})$.

D) Els tres classificadors anteriors són d'error mínim.

2 ☐ Siga un classificador en tres classes basat en les funcions discriminants lineals bidimensionals de vectors de pesos: $\mathbf{w}_1 = (0,0,1)^t$, $\mathbf{w}_2 = (0,1,0)^t$ i $\mathbf{w}_3 = (0.5,0,0)^t$. Indica quina de les figures donades a continuació és coherent amb les fronteres i regions de decisió que defineix aquest classificador.



3 ☐ Donat el classificador en dues classes definit per la frontera i regions de decisió de la figura de la dreta, quin dels següents vectors de pesos *no* defineix un classificador equivalent al donat?



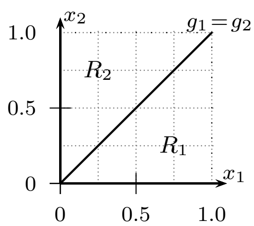
A) $\mathbf{w}_1 = (0,0,1)^t$ i $\mathbf{w}_2 = (0,1,0)^t$.

B) $\mathbf{w}_1 = (0,1,0)^t$ i $\mathbf{w}_2 = (0,0,1)^t$.

C) $\mathbf{w}_1 = (0,-1,0)^t$ i $\mathbf{w}_2 = (0,0,-1)^t$.

D) Tots els vectors de pesos anteriors defineixen classificadors equivalents.

4 ☐ Durant l'aplicació de l'algorisme Perceptró ($\alpha = 1.0$ i $b = 0$) en un problema de classificació en dues classes, s'han obtingut els vectors de pesos $\mathbf{w}_1 = (-1,1,0)^t$ i $\mathbf{w}_2 = (1,0,1)^t$. Suposa que el següent pas en l'aplicació de Perceptrón consisteix a processar una certa mostra d'entrenament $\mathbf{x}$ de classe c. Indica quina de les següents opcions donaria com a resultat un conjunt de pesos que defineix la frontera i regions de decisió de la figura de la dreta.



A) $\mathbf{x} = (-1,1)^t$ i $c = 2$.

B) $\mathbf{x} = (0,0)^t$ i $c = 2$.

C) $\mathbf{x} = (-1,1)^t$ i $c = 1$

D) $\mathbf{x} = (0,0)^t$ i $c = 1$.

5 ☐ Siga un problema de classificació en tres classes per a objectes del tipus $\mathbf{x} = (x_1, x_2)^t \in \{0,1\}^2$, amb les distribucions de probabilitat de la dreta. Quin és l'error de Bayes, ε^*, en aquest problema?

A) $\varepsilon^* < 0.2$.

B) $0.2 \le \varepsilon^* < 0.4$.

C) $0.4 \le \varepsilon^* < 0.7$.

D) $0.7 \le \varepsilon^*$.

$\mathbf{x}$		$P(c \mid \mathbf{x})$			
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

6 ☐ Es té un problema de classificació per al qual s'ha après un classificador. Així mateix, es té un conjunt de $M = 100$ mostres de test amb el qual s'ha estimat:

- La probabilitat d'error del classificador après, $\hat{p} = 0.10 = 10\,\%$.
- Un interval de confiança al $95\,\%$ per a aquesta probabilitat d'error, $\hat{I} = [0.04, 0.16] = [4\,\%, 16\,\%]$.

Es considera que la probabilitat d'error estimada és raonable i que la mateixa no variarà significativament encara que usem moltes més mostres de test. Ara bé, l'interval de confiança (al $95\,\%$) estimat, $\hat{I} = 10\,\% \pm 6\,\%$, ens sembla una mica ampli i ens preguntem si és possible reduir la seua amplitud mitjançant l'ús de més de $M = 100$ mostres de test. A més, si això fóra possible, ens preguntem si seria possible reduir aquesta amplitud a la meitat o menys; això és, tal que $\hat{I} = 10\,\% \pm \hat{R}$ amb $\hat{R} \leq 3\,\%$. En relació amb aquestes qüestions, indica quina de les següents afirmacions és correcta.

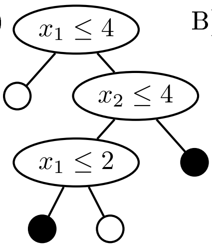
A) En general, no és possible reduir l'amplitud de $\hat{I}$ perquè $\hat{I}$ no depèn significativament de M.

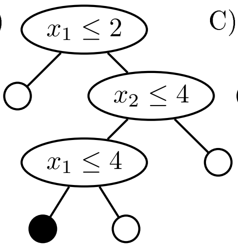
B) No és possible reduir l'amplitud de $\hat{I}$ ja que hem considerat que $\hat{p}$ no variarà significativament i, sent així, l'amplitud de $\hat{I}$ tampoc pot variar significativament.

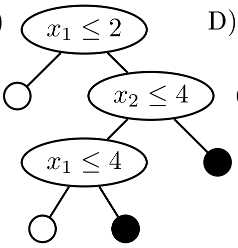
C) Sí que és possible reduir l'amplitud de $\hat{I}$, a la meitat o menys, si doblem M almenys ($M \geq 200$).

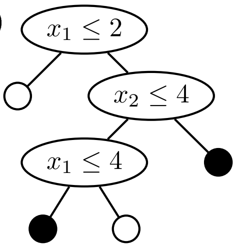
D) Sí que és possible reduir l'amplitud de $\hat{I}$, a la meitat o menys, si emprem almenys quatre vegades més mostres de test aproximadament ($M \geq 400$).

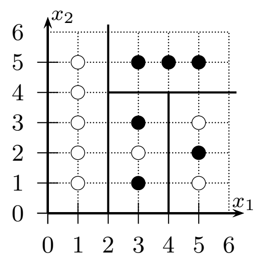
7 ☐ Donat el conjunt de mostres de 2 classes (○ i ●) de la figura de la dreta, quin dels següents arbres de classificació és coherent amb la partició representada?

A)


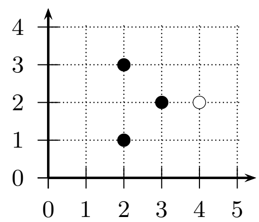
B)


C)


D)




8 ☐ La figura a la dreta mostra una partició de 4 punts bidimensionals en 2 clústers (representats mitjançant els símbols ● i ○). La transferència del punt $(3, 2)^t$ del clúster ● al clúster ○:



A) produeix un increment en la Suma d'Errors Quadràtics (SEQ).

B) produeix un decrement en la SEQ.

C) no altera la SEQ.

D) produeix una SEQ negativa.

9 ☐ En relació al càlcul de la probabilitat $P(y \mid M)$ amb la qual un model de Markov M genera una cadena de símbols y, indica quina afirmació és certa:

A) L'única manera de calcular $P(y \mid M)$ consisteix a generar explícitament totes les seqüències d'estats, calcular la probabilitat que cada seqüència d'estats haja generat y i posteriorment sumar totes les probabilitats obtingudes.

B) Una forma eficient computacionalment de calcular $P(y \mid M)$ consisteix a aplicar l'algorisme *Forward*.

C) Una forma eficient computacionalment de calcular $P(y \mid M)$ consisteix a aplicar l'algorisme de Viterbi.

D) L'única manera de calcular $P(y \mid M)$ consisteix a generar explícitament totes les seqüències d'estats mitjançant l'algorisme de Viterbi, calcular la probabilitat que cada seqüència haja generat y i sumar totes les probabilitats obtingudes.

Problema (2 punts)

Siga un model de Markov de conjunt d'estats $Q = \{1,\ 2,\ F\}$ i conjunt de símbols $\Sigma = \{a, b\}$. Es demana:

a) (1 punt) Siguen el vector de probabilitats inicials (π), matriu de transició entre estats (A) i matriu de generació de símbols (B):

π	1	2
	0.6	0.4

A	1	2	F
1	0.6	0.3	0.1
2	0.3	0.4	0.3

B	a	b
1	0.3	0.7
2	0.8	0.2

Realitza una traça de l'algorisme de Viterbi per a la cadena $y = aab$ obtenint la millor seqüència d'estats.

b) (1 punt) Siguen les tres cadenes de símbols: $y_1 = bbaa$, $y_2 = abab$ i $y_3 = aabbb$. En aplicar l'algorisme de Viterbi amb un cert model de Markov M, s'obtenen, respectivament, les següents seqüències òptimes d'estats: $1122F$, $2121F$ i $22111F$. A partir d'aquestes cadenes i les seues respectives seqüències òptimes d'estats, re-estima les probabilitats inicials (π), de transició (A) i d'emissió (B) de M (de la mateixa manera que es fa en una iteració de l'algorisme de re-estimació de Viterbi).