

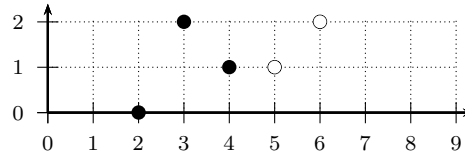
# Examen del bloc 2 de SIN: Test (1,75 punts)

ETSINF, Universitat Politècnica de València, 18 de gener de 2021

**Grup, cognoms i nom:** 3X, 3,

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

- 1 **B** La figura següent mostra una partició de 5 punts bidimensionals en dos clústers,  $\bullet$  i  $\circ$ :

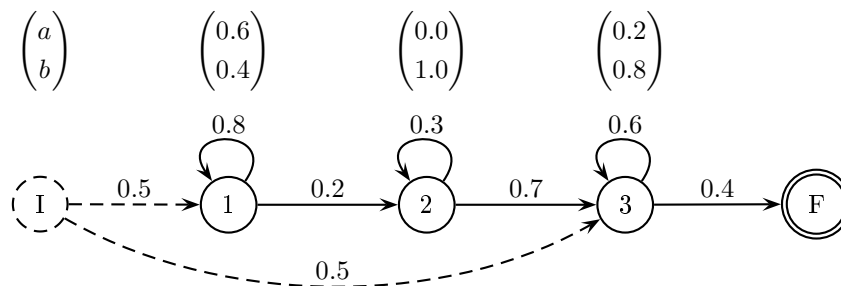


La transferència del punt  $(4, 1)^t$  del clúster  $\bullet$  al clúster  $\circ$  produeix una variació de la suma d'errors quadràtics,  $\Delta J$ , tal que:  $\Delta J = 0.166667$

- A)  $\Delta J < 0$ , açò és, la transferència és profitosa.  
B)  $0 \leq \Delta J < 1$ .  
C)  $1 \leq \Delta J < 2$ .  
D)  $\Delta J \geq 2$ .
- 2 **C** Supposeu que tenim dues caixes amb 100 pomes cadascuna. La primera caixa conté 63 pomes Gala i 37 Fuji. La segona caixa conté 50 pomes de cada tipus. Ara supposeu que s'escull una caixa a l'atzar, i després una poma a l'atzar de la caixa escollida. Si la poma escollida és Gala, la probabilitat  $P$  de que procedisca de la primera caixa és:  $P = 0.56$

- A)  $0/4 \leq P < 1/4$ .  
B)  $1/4 \leq P < 2/4$ .  
C)  $2/4 \leq P < 3/4$ .  
D)  $3/4 \leq P \leq 4/4$ .

- 3 **D** Siga  $M$  un model de Markov de representació gràfica:



¿Quantes cadenes distintes de llargària 3 pot generar  $M$ ? 8

- A) Cap.  
B) Al menys una, però no més de 3.  
C) Més de 3, però no més de 6.  
D) Més de 6.

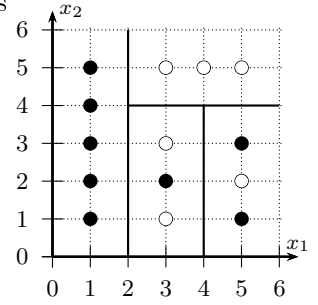
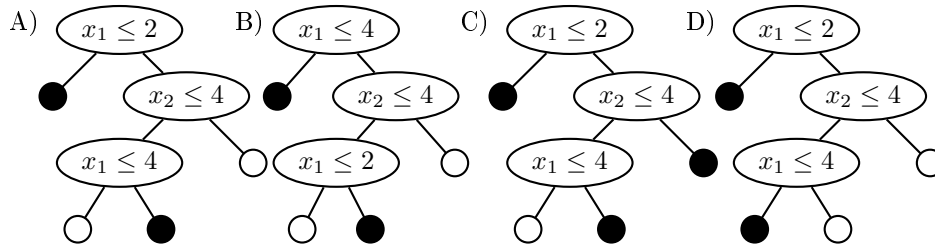
- 4 **B** Siga un problema de classificació en tres classes per a dades del tipus  $\mathbf{x} = (x_1, x_2)^t \in \{0, 1\}^2$ , amb les distribucions de probabilitat de la taula. Indica en quin interval es troba l'error de Bayes,  $\varepsilon^*$ :

$\mathbf{x}$		$P(c   \mathbf{x})$			$P(\mathbf{x})$
$x_1$	$x_2$	$c=1$	$c=2$	$c=3$	
0	0	0.2	0.7	0.1	0.4
0	1	0.1	0.5	0.4	0.1
1	0	0.2	0.3	0.5	0.5
1	1	0.1	0.6	0.3	0

$$\varepsilon^* = 0.42$$

- A)  $\varepsilon^* < 0.40$ .  
 B)  $0.40 \leq \varepsilon^* < 0.45$ .  
 C)  $0.45 \leq \varepsilon^* < 0.50$ .  
 D)  $0.50 \leq \varepsilon^*$ .

- 5 **A** Donat el conjunt de mostres de 2 classes (o i •) de la figura de la dreta, ¿quin dels següents arbres de classificació és coherent amb la partició representada?



- 6 **A** Siga  $\mathbf{x}$  un objecte a classificar en una classe de  $C$  possibles. Indica quin dels següents classificadors *no* és (de risc) d'error mínim:

- A)  $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} \log p(\mathbf{x} | c)$   
 B)  $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} \frac{p(\mathbf{x}, c)}{p(\mathbf{x})}$   
 C)  $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} p(c) p(\mathbf{x} | c)$   
 D)  $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} \log p(\mathbf{x}, c)$

7 **A** Supposeu que estem aplicant l'algorisme d'aprenentatge d'arbres de classificació per a un problema de quatre classes,  $c = 1, 2, 3, 4$ . L'algorisme ha arribat a un node  $t$  el qual inclou les següents dades: 4 de la classe 1, 4 de la 2, 2 de la 3 i 1024 de la 4. La impuresa de  $t$ ,  $\mathcal{I}(t)$ , mesurada com l'entropia de la distribució empírica de les probabilitats a posteriori de les classes en  $t$ , és:  $I = 0.09$

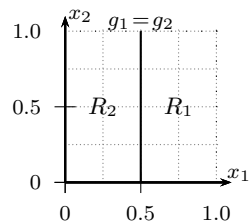
- A)  $0.00 \leq \mathcal{I}(t) < 0.50$ .
- B)  $0.50 \leq \mathcal{I}(t) < 1.00$ .
- C)  $1.00 \leq \mathcal{I}(t) < 1.50$ .
- D)  $1.50 \leq \mathcal{I}(t)$ .

8 **D** La probabilitat d'error d'un classificador s'estima que és del 17%. Determina quin és el nombre mínim de mostres de test necessari,  $M$ , per aconseguir que l'interval de confiança al 95% del dit error no supere el  $\pm 1\%$ ; açò es,  $I = [16\%, 18\%]$ :  $M = 5421$

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

9 **C** Donat el classificador en dues classes definit per la seua frontera i regions de decisió de la figura de la dreta, ¿quin dels següents vectors de pesos (en notació homogènia) defineix un classificador equivalent al donat?

- A)  $\mathbf{w}_1 = (0, -1, 0)^t$  i  $\mathbf{w}_2 = (-0.5, 0, 0)^t$ .
- B)  $\mathbf{w}_1 = (0.5, 0, 0)^t$  i  $\mathbf{w}_2 = (0, 1, 0)^t$ .
- C)  $\mathbf{w}_1 = (0, 1, 0)^t$  i  $\mathbf{w}_2 = (0.5, 0, 0)^t$ .
- D) Tots els vectors de pesos anteriors defineixen classificadors equivalents.



# Examen del bloc 2 de SIN: Problema (2 punts)

ETSINF, Universitat Politècnica de València, 18 de gener de 2021

**Grup, cognoms i nom:** 3X, 3,

## Problema sobre Viterbi

Siga  $M$  un model de Markov de conjunt d'estats  $Q = \{1, 2, F\}$ ; alfabet  $\Sigma = \{a, b\}$ ; probabilitats inicials  $\pi_1 = \frac{1}{2}, \pi_2 = \frac{1}{2}$ ; i probabilitats de transició entre estats i d'emissió de símbols:

$A$	1	2	$F$
1	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$
2	$\frac{2}{8}$	$\frac{3}{8}$	$\frac{3}{8}$

$B$	$a$	$b$
1	$\frac{2}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{2}{3}$

Es demana:

- (1 punt) Realitzeu una traça de l'algorisme de *Viterbi* per a obtindre la seqüència d'estats més probable amb la qual  $M$  genera la cadena **baa**.
- (1 punt) A partir de les cadenes d'entrenament **baa** i **bbab**, reestimeu els paràmetres d' $M$  mitjançant una iteració de l'algorisme de reestimació per Viterbi.

Solució:

- Traça de Viterbi per a la cadena **baa** (els estats 1 i 2 es representen com 0 i 1, respectivament):

	b	a	a
0	0.166692	0.055573	0.014821
1	0.333389	0.041680	0.005211
Q:	1	0	0

- Reestimació per Viterbi a partir de **baa** i **bbab**.

Per a la primera iteració, ja tenim el parell (**baa**, 211F) calculat en l'apartat anterior. Falta calcular el camí més probable per a la segona cadena d'entrenament:

	b	b	a	b
0	0.166692	0.027786	0.013896	0.001853
1	0.333389	0.083361	0.010422	0.002606
Q:	1	1	1	1

Així doncs, el segon parell és (**bbab**, 2222F). A partir d'ambdós parells, obtenim els paràmetres reestimats desitjats:

$\pi$	1	2
	$\frac{0}{2}$	$\frac{2}{2}$

$A$	1	2	$F$
1	$\frac{1}{2}$	$\frac{0}{2}$	$\frac{1}{2}$
2	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

$B$	$a$	$b$
1	$\frac{2}{2}$	$\frac{0}{2}$
2	$\frac{1}{5}$	$\frac{4}{5}$

Es pot comprovar, mitjançant una nova iteració de reestimació per Viterbi, que l'algorisme convergeix al modelo anterior.