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

ETSINF, Universitat Politècnica de València, 13 de gener de 2022

Grup, cognoms i nom: 1,

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

1 $\boxed{\text{D}}$ Siguen els següents 3 nodes d'un arbre de classificació amb mostres pertanyents a 3 classes:

c	n_1	n_2	n_3
1	4	5	5
2	5	3	4
3	4	1	2

on cada fila indica el nombre de mostres de cada classe en el node. Quina de les següents desigualtats és certa?

A) $\mathcal{I}(n_3) < \mathcal{I}(n_2) < \mathcal{I}(n_1)$

B) $\mathcal{I}(n_2) < \mathcal{I}(n_1) < \mathcal{I}(n_3)$

C) $\mathcal{I}(n_3) < \mathcal{I}(n_1) < \mathcal{I}(n_2)$

D) $\mathcal{I}(n_2) < \mathcal{I}(n_3) < \mathcal{I}(n_1)$

2 $\boxed{\text{D}}$ La probabilitat d'error d'un classificador s'estima que és del 20%. 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, $I = [19\%, 21\%]$: $M = 6147$

A) $M < 2000$.

B) $2000 \leq M < 3500$.

C) $3500 \leq M < 5000$.

D) $M \geq 5000$.

3 $\boxed{\text{C}}$ Suposeu que estem aplicant l'algorisme Perceptró, amb factor d'aprenentatge $\alpha = 1$ i marge $b = 0.1$, a un conjunt de 3 mostres bidimensionals d'aprenentatge per a un problema de 2 classes. Se sap que, després de processar les primeres 2 mostres, s'han obtés els vectors de pesos $\mathbf{w}_1 = (0, 0, 1)^t$, $\mathbf{w}_2 = (0, 0, -1)^t$. Així mateix, se sap que, després de processar l'última mostra, $(\mathbf{x}_3, c_3)$, s'obtenen els vectors de pesos $\mathbf{w}_1 = (-1, -5, -2)^t$, $\mathbf{w}_2 = (1, 5, 2)^t$. Quina de les següents mostres és eixa última mostra?

A) $((2, 4)^t, 2)$

B) $((5, 4)^t, 2)$

C) $((5, 3)^t, 2)$

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

4 $\boxed{\text{D}}$ Donat el classificador en 2 classes definit pels seus vectors de pesos $\mathbf{w}_1 = (-3, 3)^t$, $\mathbf{w}_2 = (-3, -3)^t$ en notació homogènia, quin dels següents conjunts de vectors **no** definix un classificador equivalent al donat?

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

5 $\boxed{\text{D}}$ Siga el següent conjunt de dades utilitzat per a entrenar un arbre de classificació amb 5 mostres bidimensionals que pertanyen a 2 classes:

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

Quantes particions diferents es podrien generar en el node arrel? No consideres les particions en les quals totes les dades s'assignen al mateix node fill.

A) 7
B) 8
C) 5
D) 6

6 $\boxed{\text{C}}$ Siga M un model de Markov de conjunt d'estats $Q = \{1, 2, F\}$ i alfabet $\Sigma = \{a, b\}$. Durant l'aplicació d'una iteració de l'algorisme de reestimació per Viterbi, s'ha obtés un parell "(cadena, camí més probable)" per cada cadena d'entrenament. Seguidament, a partir de tots els parells obtinguts, s'han obtingut els comptes (freqüències absolutes) de transició entre estats mostrats en la taula a la dreta. La normalització *correcta* d'aquests comptes resultarà en la taula de probabilitats de transició enntre estats:

A	1	2	F
1	4	1	4
2	1	3	2

A)

A	1	2	F
1	$\frac{4}{5}$	$\frac{1}{4}$	$\frac{4}{6}$
2	$\frac{1}{5}$	$\frac{3}{4}$	$\frac{2}{6}$

B)

A	1	2	F
1	$\frac{4}{15}$	$\frac{1}{15}$	$\frac{4}{15}$
2	$\frac{1}{15}$	$\frac{3}{15}$	$\frac{2}{15}$

C)

A	1	2	F
1	$\frac{4}{9}$	$\frac{1}{9}$	$\frac{4}{9}$
2	$\frac{1}{6}$	$\frac{3}{6}$	$\frac{2}{6}$

D)

A	1	2	F
1	$\frac{4}{4}$	$\frac{1}{4}$	$\frac{4}{4}$
2	$\frac{1}{4}$	$\frac{3}{4}$	$\frac{2}{4}$

7 $\boxed{\text{A}}$ En un problema de raonament probabilístic corresponent a desplaçaments per carretera, amb les variables aleatòries d'interés: Climatologia (C):{clar(CLA), ennuvolat (NUV), pluja (PLU)}; Lluminositat (L):{dia (DIA), nit (NIT)}; Seguretat (S):{segur (SEG), accident (ACC)}. La probabilitat conjunta de les tres variables ve donada en la taula:

	DIA			NIT		
$P(s, l, c)$	CLA	NUV	PLU	CLA	NUV	PLU
SEG	0.27	0.23	0.07	0.16	0.07	0.06
ACC	0.02	0.01	0.02	0.02	0.03	0.04

La probabilitat condicional $P(S = \text{ACC} \mid L = \text{DIA}, C = \text{NUV})$ és:

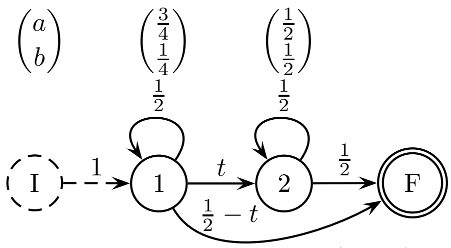
A) 0.042

B) 0.010

C) 0.240

D) 0.140

8 $\boxed{\text{C}}$ Siga M el model de Markov representat a la dreta, on t, $0 < t < \frac{1}{4}$, denota la probabilitat de transició de l'estat 1 al 2. Donada la cadena $x = \texttt{abb}$, la probabilitat de generar x mitjançant el camí $122F$, $P(\texttt{abb}, 122F)$, depén de t. Anàlogament, la probabilitat de generar x mitjançant el camí $111F$, $P(\texttt{abb}, 111F)$, també depén de t (a través de la probabilitat de transició de l'estat 1 al F). Indica en quin cas $P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F)$:



$$P(\texttt{abb}, 111F) = 1\frac{3}{4}\frac{1}{2}\frac{1}{4}\frac{1}{2}\frac{1}{4}\left(\frac{1}{2} - t\right)$$

$$P(\texttt{abb}, 122F) = 1\frac{3}{4}t\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$$

$$P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F) \rightarrow t < \tfrac{1}{10}$$

A) Mai.

B) Si i només si $0 < t < \frac{1}{20}$.

C) Si i només si $0 < t < \frac{1}{10}$.

D) Sempre, és a dir, $0 < t < \frac{1}{4}$.

9 $\boxed{\text{D}}$ Es té una partició d'un conjunt de dades 3-dimensionals en un nombre de clústers donat, $C \geq 2$. Considereu la transferència de la dada $\mathbf{x} = (1, 3, 9)^t$ d'un clúster i a altre j, $j \neq i$. Se sap que el clúster i conté 2 dades (comptant $\mathbf{x}$) i el j 3. Així mateix, se sap que la mitjana del clúster i és $\mathbf{m}_i = (4, 8, 6)^t$ i la del j $\mathbf{m}_j = (8, 10, 3)^t$. Si es realitza la dita transferència, es produirà un increment de la suma d'errors quadràtics, ΔJ, tal que: $\Delta J = 14.5$

A) $\Delta J < -70$

B) $-70 \leq \Delta J < -30$

C) $-30 \leq \Delta J < 0$

D) $\Delta J \geq 0$

Examen del bloc 2 de SIN: Problemes (2 punts)

ETSINF, Universitat Politècnica de València, 13 de gener de 2022

Grup, cognoms i nom: 1,

Problema sobre Forward i Viterbi

Siga M un model de Markov de conjunt d'estats $Q = \{1, 2, F\}$; alfabet $\Sigma = \{a, b\}$; probabilitats inicials $\pi_1 = \frac{2}{4}, \pi_2 = \frac{2}{4}$; i probabilitats de transició entre estats i d'emissió de símbols:

A	1	2	F
1	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{5}$
2	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{3}{6}$

B	a	b
1	$\frac{1}{4}$	$\frac{3}{4}$
2	$\frac{1}{2}$	$\frac{1}{2}$

Siga x=`bb`. Es demana:

1. (0,75 punts) Realitzeu una traça de l'algorisme *Forward* per a obtindre la probabilitat amb la qual M genera la cadena x, $P_M(x)$.
2. (0,75 punts) Realitzeu una traça de l'algorisme de *Viterbi* per a obtindre l'aproximació de Viterbi a la probabilitat amb la qual M genera la cadena x, $\tilde{P}_M(x)$.
3. (0,25 punts) A partir de la traça realitzada en l'apartat anterior, determineu un camí més probable amb el qual M genera x.
4. (0,25 punts) Determineu la probabilitat amb la qual M genera x seguint un camí distint al més probable determinat en l'apartat anterior.

Solució:

1. *Forward:* $P_M(x) = 199/2400 = 0.08292$

	b	b	
1	3/8	7/40	
2	1/4	23/240	
F			199/2400

2. Viterbi: $\tilde{P}_M(x) = 3/80 = 0.03750$

	b	b	
1	3/8	9/80	
2	1/4	3/40	
F			3/80

3. Camí més probable: $12F$
4. $P_M(x) - \tilde{P}_M(x) = 109/2400 = 0.04542$.