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

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

Grup, cognoms i nom: 2,

Marca cada requadre amb una única opció. Puntuació: max(0, (encerts − errors / 3) · 1,75 / 9).

1 $\boxed{\text{A}}$ La probabilitat d'error d'un classificador s'estima que és del 2%. 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 ±1%; açò es, $I = [1\%, 3\%]$: $M = 753$

A) $M < 2000$.

B) $2000 \leq M < 3500$.

C) $3500 \leq M < 5000$.

D) $M \geq 5000$.

2 $\boxed{\text{B}}$ 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 $\mathbf{x} = (6, 1, 10)^t$ d'un clúster i a altre j, $j \neq i$. Se sap que el clúster i conté 4 dades (comptant $\mathbf{x}$) i el j 2. Així mateix, se sap que la mitjana del clúster i és $\mathbf{m}_i = (2, 3, 3)^t$ i la del j $\mathbf{m}_j = (7, 10, 9)^t$. Si es realitza la dita transferència, es produirà un increment de la suma d'errors quadràtics, ΔJ, tal que: $\Delta J = -36.7$

A) $\Delta J < -70$

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

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

D) $\Delta J \geq 0$

3 $\boxed{\text{A}}$ 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}	4	2	1	2	2
x_{n2}	1	2	2	5	4
c_n	2	2	1	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) 5

B) 3

C) 7

D) 4

4 $\boxed{B}$ 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	1	3	4
2	4	1	1

A)

A	1	2	F
1	$\frac{1}{14}$	$\frac{3}{14}$	$\frac{4}{14}$
2	$\frac{4}{14}$	$\frac{1}{14}$	$\frac{1}{14}$

B)

A	1	2	F
1	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$
2	$\frac{4}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

C)

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

D)

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

5 $\boxed{A}$ Siguen els següents 3 nodes d'un arbre de classificació amb mostres pertanyents a 3 classes:

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

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

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

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

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

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

6 $\boxed{A}$ 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 $\mathbf{w}_1 = (0, -1, -2)^t$, $\mathbf{w}_2 = (0, 1, 2)^t$. Així mateix, se sap que, després de processar l'última mostra, $(\mathbf{x}_3, c_3)$, s'obtenen els vectors de pesos $\mathbf{w}_1 = (1, 4, 0)^t$, $\mathbf{w}_2 = (-1, -4, 0)^t$. Quina de les següents mostres és eixa última mostra?

A) $((5, 2)^t, 1)$

B) $((4, 1)^t, 2)$

C) $((1, 2)^t, 2)$

D) $((4, 3)^t, 2)$

7 $\boxed{\text{D}}$ 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	0.33	0.23	0.04	0.10	0.07	0.09
ACC	0.03	0.02	0.01	0.01	0.02	0.05

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

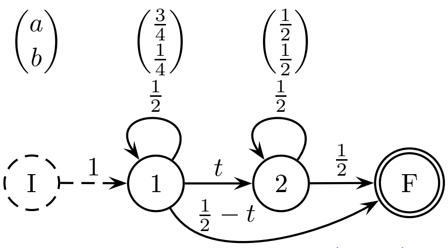
A) 0.140
B) 0.360
C) 0.030
D) 0.083

8 $\boxed{\text{B}}$ Donat el classificador en 2 classes definit pels seus vectors de pesos $\mathbf{w}_1 = (-1, 0, 2)^t$, $\mathbf{w}_2 = (-1, 3, -1)^t$ en notació homogènia, quin dels següents conjunts de vectors **no** definix un classificador equivalent al donat?

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

9 $\boxed{\text{C}}$ 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 $P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F)$:

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, $0 < t < \frac{1}{4}$.



$$P(\texttt{abb}, 111F) = 1\frac{3}{4}\frac{1}{2}\frac{1}{4}\frac{1}{2}\frac{1}{4}\left(\frac{1}{2} - t\right)$$
$$P(\texttt{abb}, 122F) = 1\frac{3}{4}t\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}$$
$$P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F) \to t < \frac{1}{10}$$

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

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

Grup, cognoms i nom: 2,

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 $\pi_1 = \frac{1}{2}, \pi_2 = \frac{1}{2}$; i probabilitats de transició entre estats i d'emissió de símbols:

A	1	2	F
1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{7}$	$\frac{3}{7}$	$\frac{3}{7}$

B	a	b
1	$\frac{3}{6}$	$\frac{3}{6}$
2	$\frac{1}{3}$	$\frac{2}{3}$

Siga x=`ab`. 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:* $P_M(x) = 73/1176 = 0.06207$

```
           a          b
1        1/4       3/56
2        1/6     13/126
F                            73/1176
```

2. Viterbi: $\tilde{P}_M(x) = 1/42 = 0.02381$

```
           a          b
1        1/4       1/24
2        1/6       1/18
F                               1/42
```

3. Camí més probable: $12F$
4. $P_M(x) - \tilde{P}_M(x) = 15/392 = 0.03827$.