Examen Final de Sistemes Intel·ligents: Bloc 2

ETSINF, Universitat Politècnica de València, 26 de gener de 2015

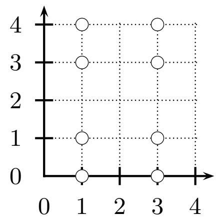
Cognoms: ____________________ **Nom:** ____________

Grup: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3E** ☐ **3F** ☐ **RE1** ☐ **RE2**

Qüestions (2 punts; temps estimat: 30 minuts)

Marca cada requadre amb una única opció d'entre les donades.

1 $\boxed{\text{B}}$ La figura a la dreta mostra 8 punts bidimensionals. La menor "Suma d'Errors Quadràtics", J, amb la qual poden agrupar-se aquests punts en dos clústers és:



A) $0 \leq J \leq 7$
B) $7 < J \leq 14$ $\quad J = 10$
C) $14 < J \leq 21$
D) $21 < J$

2 $\boxed{\text{D}}$ Siguen X, Y i Z tres variables aleatòries. Es diu que X i Y són *condicionalment independents* donada Z si i solament si

$$P(X=x, Y=y \mid Z=z) = P(X=x \mid Z=z)\, P(Y=y \mid Z=z) \qquad \text{per a tot } x,\ y \text{ i } z.$$

Si es compleix aquesta igualtat, podem calcular $P(Z=z \mid X=x, Y=y)$ com segueix:

A) $P(Z=z \mid X=x, Y=y) = \dfrac{P(X=x, Y=y, Z=z)}{P(X=x, Y=y)}$
B) $P(Z=z \mid X=x, Y=y) = \dfrac{P(Z=z)\, P(X=x, Y=y \mid Z=z)}{P(X=x, Y=y)}$
C) $P(Z=z \mid X=x, Y=y) = \dfrac{P(Z=z)\, P(X=x \mid Z=z)\, P(Y=y \mid Z=z)}{P(X=x, Y=y)}$
D) De les tres maneres anteriors.

3 $\boxed{\text{C}}$ Es vol construir un sistema de reconeixement de formes per a dígits manuscrits representats mitjançant cadenes de contorn de 4 direccions; açò és, mitjançant cadenes de símbols en l'alfabet $\Sigma = \{1, 2, 3, 4\}$. Donada una seqüència de cadenes d'entrenament amb les seues corresponents etiquetes de classe, construirem el sistema com segueix:

A) Emprarem l'algorisme Perceptró i obtindrem un classificador lineal.
B) Aprendrem un Arbre de Decisió i Classificació mitjançant l'algorisme ADC.
C) Dissenyarem un classificador basat en models de Markov aplicant l'algorisme de re-estimació per Viterbi.
D) Les tres opcions anteriors són vàlides.

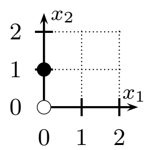
4 $\boxed{\text{C}}$ En un problema de classificació en tres classes ($C = \{a, b, c\}$), en el qual es disposa de 100 mostres de la classe a, 100 mostres de la classe b i 100 mostres de la classe c, siga y un fet o dada. La decisió òptima de classificació per a y és la classe a amb una probabilitat a posteriori de 0.50. Quina de les següents afirmacions és correcta?

A) $P(C=a \mid Y=y) > P(C=b \mid Y=y) + P(C=c \mid Y=y)$
B) $P(Y=y \mid C=a) = \dfrac{0.5\ P(C=a)}{P(Y=y)}$
C) $P(Y=y \mid C=a) = P(Y=y \mid C=b) + P(Y=y \mid C=c)$
D) Cap de les anteriors.

5 $\boxed{\text{C}}$ Donat un classificador lineal de 2 classes $\circ$ i $\bullet$ definit pel seu conjunt de pesos $\mathbf{a}_\circ = (0, -1, 1)^t$ i $\mathbf{a}_\bullet = (0, 1, -1)^t$, Què conjunt de pesos dels següents no defineix un classificador equivalent al donat?

A) $\mathbf{a}_\circ = (1, -1, 1)^t$ i $\mathbf{a}_\bullet = (1, 1, -1)^t$ $\quad f(z) = az + b$ amb $a = 1$ i $b = 1$
B) $\mathbf{a}_\circ = (-1, -2, 2)^t$ i $\mathbf{a}_\bullet = (-1, 2, -2)^t$ $\quad f(z) = az + b$ amb $a = 2$ i $b = -1$
C) $\mathbf{a}_\circ = (0, 2, -2)^t$ i $\mathbf{a}_\bullet = (0, -2, 2)^t$ $\quad f(z) = az + b$ amb $a = -2$ i $b = 0$
D) $\mathbf{a}_\circ = (0, -2, 2)^t$ i $\mathbf{a}_\bullet = (0, 2, -2)^t$ $\quad f(z) = az + b$ amb $a = 2$ i $b = 0$

6 $\boxed{\text{A}}$ En la figura de la dreta es representen dues mostres d'aprenentatge bidimensionals de 2 classes: $(x_1, \circ)$ i $(x_2, \bullet)$. Donats el conjunt de pesos $\mathbf{a}_\circ = (0, 1, -2)^t$ i $\mathbf{a}_\bullet = (0, 0, 1)^t$, si apliquem una iteració de l'algorisme Perceptró amb factor d'aprenentatge $\alpha = 1.0$ i marge $b = 0.5$ a partir del conjunt de pesos i mostres d'aprenentatge donats, quants errors de classificació es produeixen sobre les mostres d'aprenentatge amb el nou conjunt de pesos?



A) 0 $\quad \mathbf{a}_\circ = (1, 1, -2)^t$ i $\mathbf{a}_\bullet = (-1, 0, 1)^t$
B) 1
C) 2
D) 3

Examen Final de Sistemas Intel·ligents: Bloc 2

ETSINF, Universitat Politècnica de València, 26 de gener de 2015

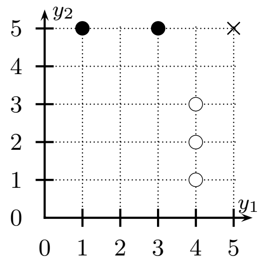
Cognoms: **Nom:**

Grup: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3E** ☐ **3F** ☐ **RE1** ☐ **RE2**

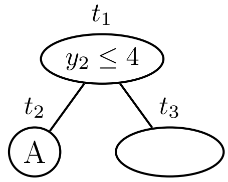
Problemes (3 punts; temps estimat: 60 minuts)

1. **(1.5 punts)**

Per a aprendre un arbre de classificació es disposa d'una mostra d'entrenament formada per 6 vectors bidimensionals pertanyents a 3 classes, A, B i C. Aquests vectors es mostren en la figura a la dreta ($A = \circ$, $B = \bullet$ i $C = \times$). En les primeres invocacions recursives de l'algorisme ADC (amb $\epsilon = 0.5$ bits) s'ha produït el sub-arbre amb tres nodes que es mostra en la figura de baix. Aquest sub-arbre correspon a una primera divisió òptima de la mostra d'entrenament en dos subconjunts mitjançant l'*"split"* $(2, 4.0)$ (és a dir, $y_2 \leq 4$). En aquest procés inicial s'han obtingut els paràmetres que es mostren en la taula.



Node	Split	$P(A \mid t_i)$	$P(B \mid t_i)$	$P(C \mid t_i)$	$P_{t_i}(L)$	$P_{t_i}(R)$	$\mathcal{I}(t_i)$	$\Delta\mathcal{I}(t_1)$
t_1	(2,4)	1/2	1/3	1/6	1/2	1/2	1.459	1.000
t_2	–	1	0	0	–	–	0	–
t_3		0	2/3	1/3				
t_4								
t_5								



(a) Expliqueu com s'obtenen els següents valors de la taula: $P(A \mid t_1)$, $P(B \mid t_1)$, $P(C \mid t_1)$, $P_{t_1}(R)$ i $\mathcal{I}(t_1)$.

El node arrel (t_1) representa als 6 les dades disponibles. D'ells hi ha 3 de la classe A, 2 de la classe B i 1 de la classe C. Per tant: $P(A \mid t_1) = 3/6 = 1/2$, $\quad P(B \mid t_1) = 2/6 = 1/3$, $\quad P(C \mid t_1) = 1/6$

L'*"split"* $(2, 4.0)$ ($t_2 \leq 4$) divideix l'arbre arrel en dos subárbres: un arrelat en t_2, que representa 3 dades tals que $y_2 \leq 4$, i un altre en t_3, que representa altres 3 dades tals que $y_2 > 4$. Així doncs: $P_{t_1}(R) = 3/6 = 1/2$

Finalment: $\mathcal{I}(t_1) \;=\; -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{3}\log_2\frac{1}{3} - \frac{1}{6}\log_2\frac{1}{6} \;\approx\; 1.459$ bits

(b) Calculeu la impuresa del node t_3.

$\mathcal{I}(t_3) \;=\; 0 - \frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} \;\approx\; 0.918$ bits

(c) Trobeu l'*"split"* òptim per al node t_3, completeu l'execució de l'algorisme ADC i completeu les cel·les de taula que estan en blanc.

El node t_3 representa els vectors $((1,5)^t, B), ((3,5)^t, B), ((5,5)^t, C)$ per als quals solament hi ha dues particions possibles, corresponents als *"splits"*: $y_1 \leq 2$ i $y_1 \leq 4$. Els decrements d'impuresa corresponents són:

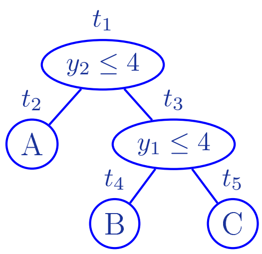
$\Delta\mathcal{I}(1,2,t_3) = \mathcal{I}(t_3) - \frac{1}{3}\mathcal{I}(t_4) - \frac{2}{3}\mathcal{I}(t_5) \;\approx\; 0.918 - 0 - \frac{2}{3}\cdot 1 \;=\; 0.251$ bits

$\Delta\mathcal{I}(1,4,t_3) = \mathcal{I}(t_3) - \frac{2}{3}\mathcal{I}(t_4) - \frac{1}{3}\mathcal{I}(t_5) \;\approx\; 0.918 - 0 - 0 \;=\; 0.918$ bits

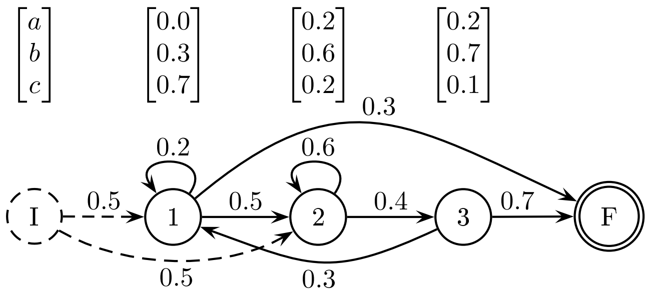
D'aquests, el major decrement és per a l'split $(1, 4)$ (o siga, $y_1 \leq 4$).

L'arbre resultant i els paràmetres corresponents es mostren baix en la figura i taula, respectivament.

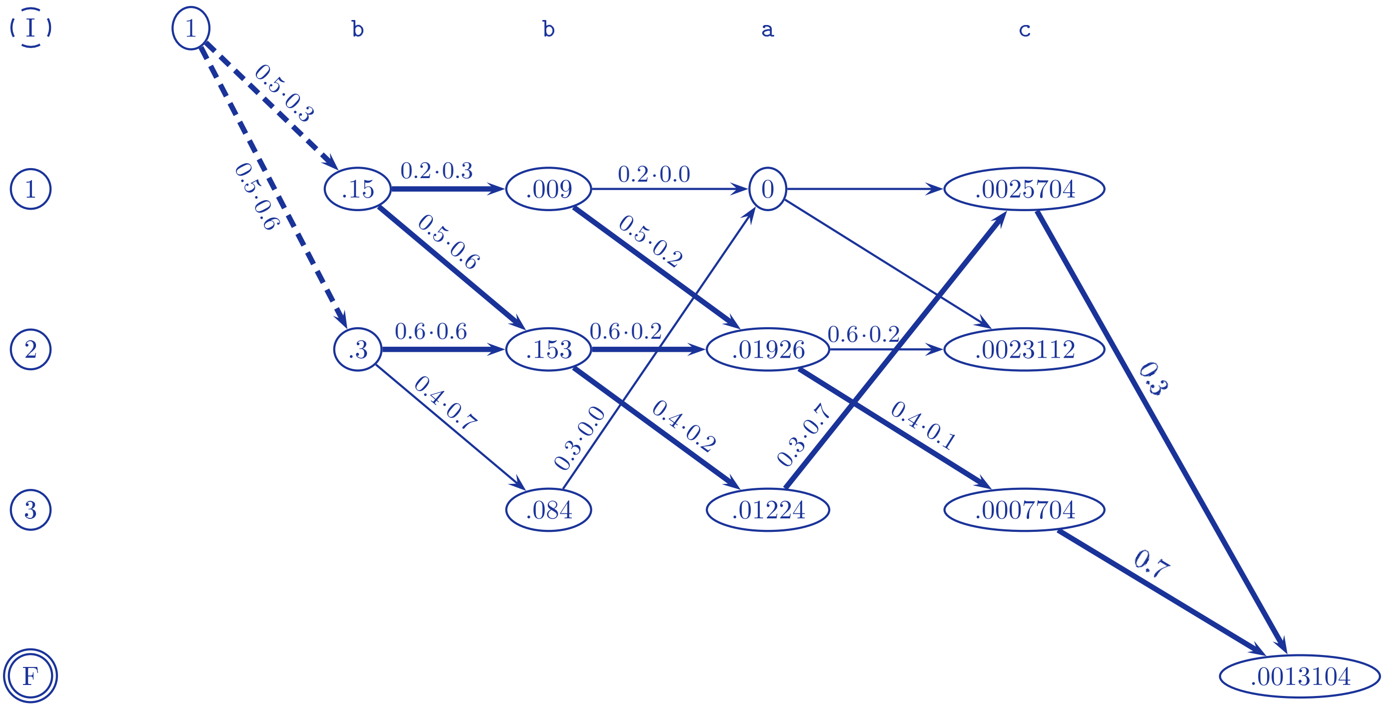
Node	Split	$P(A \mid t_i)$	$P(B \mid t_i)$	$P(C \mid t_i)$	$P_{t_i}(L)$	$P_{t_i}(R)$	$\mathcal{I}(t_i)$	$\Delta\mathcal{I}(t_1)$
t_1	(2,4)	1/2	1/3	1/6	1/2	1/2	1.459	1.000
t_2	–	1	0	0	–	–	0	–
t_3	(1,4)	0	2/3	1/3	2/3	1/3	0.918	0.918
t_4	–	0	1	0	–	–	0	–
t_5	–	0	0	1	–	–	0	–



2. **(1.5 punts)** Siga M el model de Markov:



Calculeu la probabilitat exacta de que M genere la cadena *bbac*, $P_M(bbac)$, mitjançant l'algorisme *Forward*.



$$P_M(bbac) \; = \; 0.0013104$$