Examen del Bloc 2 de Sistemes Intel·ligents

ETSINF, Universitat Politècnica de València, 15 de gener de 2014

Cognoms: **Nom:**

Grup: ☐ **3A** ☐ **3B** ☐ **3C** ☐ **3D** ☐ **3I** ☐ **RE1** ☐ **RE2**

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

1 $\boxed{\text{B}}$ Donada la probabilitat conjunta de dues variables aleatòries X i Y, la probabilitat condicional $P(Y = y \mid X = x)$ es pot obtenir mitjançant:

A) $P(y \mid x) \;=\; 1 \;/\; P(x, y)$
B) $P(y \mid x) \;=\; P(x, y) \;/\; \sum_{y'} P(x, y')$
C) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \;/\; \sum_{y'} P(x, y')$
D) $P(y \mid x) \;=\; \sum_{x'} P(x', y) \cdot \sum_{y'} P(x, y')$

2 $\boxed{\text{A}}$ En un problema de decisió binari ($D = \{0, 1\}$), siga y un fet o dada i $d^\star(y) = 0$ la decisió de mínim error per a aqueix y. Identifica quina de les següents expressions determina *incorrectament* la mínima probabilitat d'error per a y:

A) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 1 \mid Y = y)$
B) $P_\star(\text{error} \mid Y = y) \;=\; 1 - P(D = 0 \mid Y = y)$
C) $P_\star(\text{error} \mid Y = y) \;=\; P(D = 1 \mid Y = y)$
D) $P_\star(\text{error} \mid Y = y) \;=\; 1 - \text{máx}_d\, P(D = d \mid Y = y)$

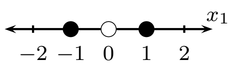
3 $\boxed{\text{D}}$ En un problema de diagnòstic diferencial entre *Grip* i *Refredat*, se sap que la incidència relativa de la *Grip* respecte al *Refredat* és del 30 % i es coneixen les següents distribucions de temperatures corporals:

$t(^oC)$	36	37	38	39	40
$P(T = t \mid D = \text{GRIP})$	0.05	0.10	0.20	0.30	0.35
$P(T = t \mid D = \text{REFR})$	0.10	0.30	0.40	0.15	0.05

La probabilitat a posteriori de que un pacient amb 38º de febre tinga *Grip* és:

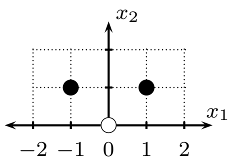
A) major que 0.8
B) menor que 0.1
C) entre 0.3 i 0.6
D) menor que la probabilitat que amb aqueixa temperatura tinga *Refredat*

4 $\boxed{\text{B}}$ En la figura de la dreta es representen tres mostres d'aprenentatge unidimensionals de 2 classes: ○ i ●. Quin serà el nombre d'errors de classificació comesos per un classificador lineal de mínim error?



A) 0
B) 1
C) 2
D) 3

5 $\boxed{\text{A}}$ Suposem que en l'exercici anterior afegim un nova característica x_2 que es defineix com $x_2 = x_1^2$. D'aquesta forma les tres mostres d'aprenentatge passen a ser bidimensionals com s'observa en la figura de la dreta. En aquest cas, quin serà el nombre d'errors de classificació comesos per un classificador lineal de mínim error?



A) 0
B) 1
C) 2
D) 3

6 $\boxed{A}$ Siga un problema de classificació en 2 classes, $c = 1, 2$, per a objectes representats mitjançant vectors de característiques bidimensionals. Es tenen 2 mostres d'entrenament: $\mathbf{x}_1 = (0, 0)^t$ de la classe $c_1 = 1$, i $\mathbf{x}_2 = (1, 1)^t$ de $c_2 = 2$. Així mateix, es té un classificador lineal definit pels vectors de pesos: $\mathbf{w}_1 = (w_{10}, w_{11}, w_{12}) = (1, -1, -1)^t$ i $\mathbf{w}_2 = (w_{20}, w_{21}, w_{22}) = (-1, 1, 1)^t$. Si apliquem una iteració de l'algorisme Perceptró a partir d'aquests vectors de pesos, amb factor d'aprenentatge $\alpha = 1$ i marge $b = 0.1$, llavors:

A) No es modificarà cap vector de pesos.
B) Es modificarà el vector de pesos de la classe 1.
C) Es modificarà el vector de pesos de la classe 2.
D) Es modificaran els vectors de pesos d'ambdues classes.

7 $\boxed{C}$ L'algorisme Perceptró està governat per dos paràmetres que anomenem *velocitat d'aprenentatge*, α, i *marge*, b, sent tots dos valors reals. En cas que no sabérem si les mostres d'aprenentatge són linealment separables, quins valors dels paràmetres α i b proporcionen majors garanties d'obtenir fronteres de decisió de millor qualitat?

A) $\alpha = 0.1$ i $b = 0.0$.
B) $\alpha = 0.0$ i $b = 0.0$.
C) $\alpha = 0.1$ i $b = 1.0$.
D) $\alpha = 0.0$ i $b = 1.0$.

8 $\boxed{C}$ Siga un problema de classificació en C classes, $c = 1, \dots, C$, per al qual s'ha après un arbre de classificació T. Siga t un node de T la impuresa del qual ve donada mitjançant l'entropia, $H(t)$, associada a les probabilitats a posteriori de les classes en t, $P(1 \mid t), \dots, P(C \mid t)$. El node t serà màximament pur quan:

A) Les classes siguen equiprobables; açò és, $P(1 \mid t) = \cdots = P(C \mid t) = \frac{1}{C}$.
B) Existisca una classe c^* de major probabilitat que la resta; açò és, $P(c^* \mid t) > P(c \mid t)$ per a tot $c \neq c^*$.
C) Existisca una classe c^* de probabilitat 1; açò és, tal que $P(c^* \mid t) = 1$.
D) Cap de les anteriors.

9 $\boxed{A}$ Siga un problema de classificació en 2 classes, $c = 1, 2$, per a objectes representats mitjançant vectors de característiques reals bidimensionals; açò és, de la forma $\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$. Es tenen 4 mostres d'entrenament: $\mathbf{y_1} = (1, 0.2)^t$, pertanyent a la classe 1; i $\mathbf{y_2} = (2, 0.2)^t$, $\mathbf{y_3} = (3, 0.8)^t$ i $\mathbf{y_4} = (1, 0.8)^t$, pertanyents a la classe 2. Volem construir un arbre de classificació emprant el decrement d'impuresa (mesurat en termes d'entropia) per a mesurar la qualitat d'una partició d'un node. En el cas del node arrel i considerant només la característica y_1, quina de les següents afirmacions és *certa*? (Nota: $\log_2(1/3) = -1.585$ i $\log_2(2/3) = -0.585$).

A) La millor partició és $y_1 \leq 1$.
B) La millor partició és $y_1 \leq 2$.
C) La millor partició és $y_1 \leq 3$.
D) Cap de les anteriors.

10 $\boxed{C}$ Siga un problema de classificació en C classes, $c = 1, \dots, C$, per al qual s'ha après un arbre de classificació T. Siga t un node terminal de T en el qual s'han estimat les probabilitats a posteriori de les classes $\hat{P}(1 \mid t), \dots, \hat{P}(C \mid t)$. Un criteri simple i eficaç per a assignar una etiqueta de classe a t és:

A) La d'una classe de probabilitat a posteriori mínima.
B) La d'una classe de probabilitat a posteriori pròxima a la mitjana (i.e. $\frac{1}{C}$).
C) La d'una classe de probabilitat a posteriori màxima.
D) Cap de les anteriors.

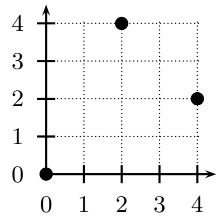
11 $\boxed{D}$ Indica quina de la següents afirmacions sobre *Clustering* és correcta:

A) Se sol emprar l'algorisme *Perceptró* a partir de dades d'entrenament *amb* etiquetes de classe.
B) Se sol emprar l'algorisme *Perceptró* a partir de dades d'entrenament *sense* etiquetes de classe.
C) Se sol emprar l'algorisme *C-Mitjanes* a partir de dades d'entrenament *amb* etiquetes de classe.
D) Se sol emprar l'algorisme *C-Mitjanes* a partir de dades d'entrenament *sense* etiquetes de classe.

12 $\boxed{D}$ El criteri de clustering particional "Suma d'Errors Quadràtics" és apropiat quan les dades formen clústers:

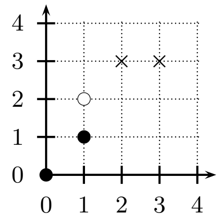
A) No allargats.
B) Allargats i de qualsevol grandària.
C) Allargats i de grandària similar.
D) Cap de les anteriors.

13 $\boxed{\text{B}}$ La menor suma d'errors quadràtics amb la qual poden agrupar-se en dos clústers els punts a la dreta és un valor:



A) Entre 0 i 3.
B) Entre 3 i 6. $J = 4$
C) Entre 6 i 9.
D) Major que 9.

14 $\boxed{\text{B}}$ La figura a la dreta mostra una partició de 5 punts bidimensionals en 3 clústers (representats mitjançant els símbols $\bullet$, $\circ$ i $\times$). Considera totes les possibles transferències de clúster de cada punt (en un clúster no unitari). En termes de suma d'errors quadràtics (J):

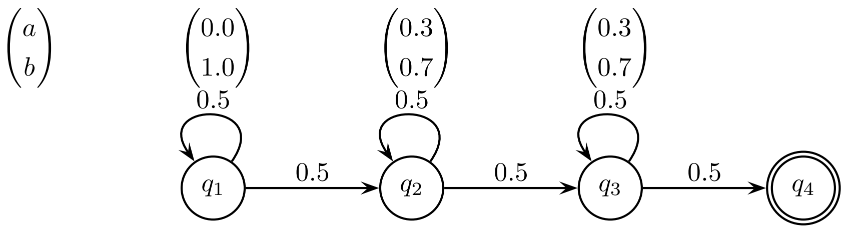


A) Cap transferència permet millorar J.
B) Només es pot millorar J transferint $(1,1)^t$ del clúster $\bullet$ al $\circ$.
C) Només es pot millorar J transferint $(2,3)^t$ del clúster $\times$ al $\circ$.
D) Les dues transferències anteriors permeten millorar J.

15 $\boxed{\text{C}}$ Donat un Model Ocult de Markov Θ i una cadena y reconeguda per aquest model, indica quina de les següents afirmacions és certa:

A) Sempre es compleix que $P(y|\Theta) = \widetilde{P}(y|\Theta)$.
B) Sempre es compleix que $P(y|\Theta) \leq \widetilde{P}(y|\Theta)$.
C) Sempre es compleix que $P(y|\Theta) \geq \widetilde{P}(y|\Theta)$.
D) Sempre es compleix que $P(y|\Theta) \neq \widetilde{P}(y|\Theta)$.

16 $\boxed{\text{C}}$ Donat el Model Ocult de Markov Θ



amb $\pi_{q_1} = 1, \pi_{q_2} = \pi_{q_3} = \pi_{q_4} = 0$ i les cadenes $y_1 =$ "babb" i $y_2 =$ "aaaa", indica quina de les següents afirmacions és certa:

A) $P(y_1|\Theta) = P(y_2|\Theta)$.
B) $P(y_1|\Theta) < P(y_2|\Theta)$.
C) $P(y_1|\Theta) > P(y_2|\Theta)$.
D) $P(y_1|\Theta) = P(y_2|\Theta) = 0$.

17 $\boxed{\text{D}}$ Donat el Model Ocult de Markov Θ de la pregunta anterior, si l'estimem amb una sola iteració amb la mostra $Y = \{baba, abab\}$ i utilitzant l'algorisme de Viterbi, indica quina de les següents afirmacions és certa:

A) El model estimat té tots els paràmetres a 0.0.
B) Cap probabilitat de transició entre estats en el model estimat pren valor 0.0.
C) El model obtingut té tots els paràmetres igual al model incial.
D) El model obtingut queda amb diversos paràmetres amb valor 0.

18 $\boxed{\text{C}}$ En relació a l'algorisme *forward* definit per a Models Ocults de Markov, indica quina de la següents afirmacions és vertadera:

A) Calcula la probabilitat d'una cadena tenint en compte només la seqüència d'anàlisi de màxima probabilitat.
B) Calcula la probabilitat d'una cadena sense incloure la probabilitat de la seqüència d'estats més probable.
C) Calcula la probabilitat d'una cadena incloent totes les seqüències d'estats.
D) Mai calcula la probabilitat d'una cadena.