

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

ETSINF, Universitat Politècnica de València, 16 de gener de 2024

Grup, cognoms i nom: 2,

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

- 1 ☒ En un problema de raonament probabilístic corresponent a diagnòstic de grip, les variables aleatòries d'interès són: Grip (G):{positiu (POS), negatiu (NEG)}; Ventilació (V):{alta (ALT), baixa (BAI)}; Activitat (A):{silenci (SIL), parlant (PAR), exercici (EXE)}. La probabilitat conjunta de les tres variables ve donada en la taula següent: $P = 0.56$

	ALT			BAI		
$P(g, v, a)$	SIL	PAR	EXE	SIL	PAR	EXE
POS	0.01	0.02	0.02	0.01	0.03	0.05
NEG	0.29	0.19	0.10	0.14	0.10	0.04

La probabilitat condicional $P(G = \text{POS} \mid V = \text{BAI}, A = \text{EXE})$ és:

- A) $P \leq 0.25$
 B) $0.25 < P \leq 0.50$
 C) $0.50 < P \leq 0.75$
 D) $0.75 < P \leq 1.0$
- 2 ☒ 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 (o escull l'última opció si els tres són d'error mínim):
- A) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} \log p(\mathbf{x} \mid c) + \log p(c)$
 B) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} e^{p(c|\mathbf{x})} + e^{p(\mathbf{x})}$
 C) $c(\mathbf{x}) = \arg \max_{c=1, \dots, C} e^{p(\mathbf{x}, c)} - e^{p(\mathbf{x})}$
 D) Els tres classificadors anteriors són d'error mínim.

- 3 ☒ 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 del classificador $c(\mathbf{x})$ donat en la taula, ε :

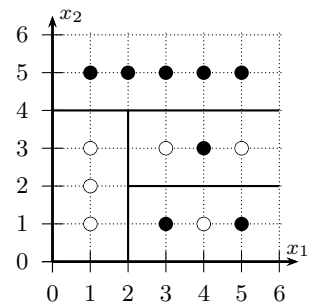
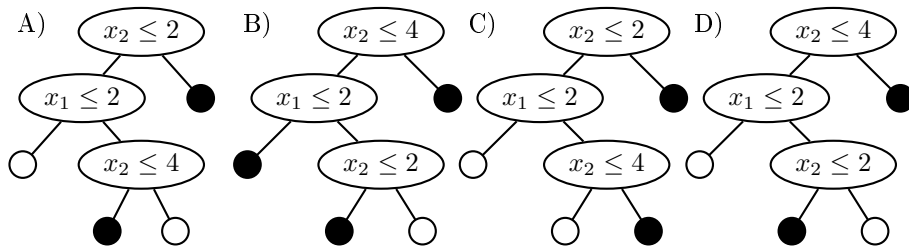
- A) $\varepsilon < 0.25$.
 B) $0.25 \leq \varepsilon < 0.50$.
 C) $0.50 \leq \varepsilon < 0.75$.
 D) $0.75 \leq \varepsilon$.

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

$\varepsilon = 0.69$

- 4 ☒ C Supposeu 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 obtingut els vectors de pesos $\mathbf{w}_1 = (0, 2, 1)^t$, $\mathbf{w}_2 = (0, -2, -1)^t$. Així mateix, se sap que, després de processar l'última mostra, (\mathbf{x}_3, c_3) , s'obtenen els vectors de pesos $\mathbf{w}_1 = (-1, 1, -3)^t$, $\mathbf{w}_2 = (1, -1, 3)^t$. Quina de les següents mostres és eixa última mostra?
- A) $((2, 1)^t, 1)$
 - B) $((3, 1)^t, 1)$
 - C) $((1, 4)^t, 2)$
 - D) $((2, 4)^t, 1)$
- 5 ☒ A Donat el classificador en 3 classes definit pels seus vectors de pesos $\mathbf{w}_1 = (2, 1, 1)^t$, $\mathbf{w}_2 = (1, -3, -3)^t$, $\mathbf{w}_3 = (2, 0, -1)^t$ en notació homogènia, quin dels següents conjunts de vectors **no** defineix un classificador equivalent al donat?
- A) $\mathbf{w}_1 = (-2, -1, -1)^t$, $\mathbf{w}_2 = (-1, 3, 3)^t$, $\mathbf{w}_3 = (-2, 0, 1)^t$
 - B) $\mathbf{w}_1 = (4, 2, 2)^t$, $\mathbf{w}_2 = (2, -6, -6)^t$, $\mathbf{w}_3 = (4, 0, -2)^t$
 - C) $\mathbf{w}_1 = (4, 1, 1)^t$, $\mathbf{w}_2 = (3, -3, -3)^t$, $\mathbf{w}_3 = (4, 0, -1)^t$
 - D) $\mathbf{w}_1 = (6, 2, 2)^t$, $\mathbf{w}_2 = (4, -6, -6)^t$, $\mathbf{w}_3 = (6, 0, -2)^t$
- 6 ☒ D Indica quina de les següents afirmacions sobre regressió logística és *incorrecta* (o escull l'última opció si les tres primeres són correctes):
- A) Regressió logística és un model probabilístic de classificació basat en una funció predictora de logits lineal amb l'entrada
 - B) Al tractar-se d'un model probabilístic de classificació, regressió logística permet aplicar regles de decisió més generals que la MAP (decidir-se per la classe de màxima probabilitat a posteriori)
 - C) Regressió logística és un model probabilístic de classificació basat en la distribució categòrica
 - D) Les tres afirmacions anteriors són correctes

- 7 **D** Donat el conjunt de mostres de 2 classes (\circ i \bullet) de la figura de la dreta, ¿quin dels següents arbres de classificació és coherent amb la partició representada?



- 8 **A** Supposeu que estem aplicant l'algorisme d'aprenentatge d'arbres de classificació per a un problema de 3 classes, $c = 1, 2, 3$. L'algorisme ha arribat a un node t que ha estat dividit en un node esquerre amb 3 mostres de la classe 1, 1 mostra de la classe 2 i 2 mostres de la classe 3; i un node dret amb 1 mostra de la classe 1, 0 mostres de la classe 2 i 0 mostres de la classe 3. Quin decrement d'impuresa s'ha aconseguit amb esta partició? $\Delta \mathcal{I} = 0.13$

- A) $0.00 \leq \Delta \mathcal{I} < 0.25$.
 B) $0.25 \leq \Delta \mathcal{I} < 0.50$.
 C) $0.50 \leq \Delta \mathcal{I} < 0.75$.
 D) $0.75 \leq \Delta \mathcal{I}$.

- 9 **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 $\mathbf{x} = (4, 10, 4)^t$ 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 $\mathbf{m}_i = (1, 8, 2)^t$ i la del j $\mathbf{m}_j = (10, 2, 10)^t$. Si es realitza la dita transferència, es produirà un increment de la suma d'errors quadràtics, ΔJ , tal que: $\Delta J = 68.0$

- A) $\Delta J < -70$
 B) $-70 \leq \Delta J < -30$
 C) $-30 \leq \Delta J < 0$
 D) $\Delta J \geq 0$

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

ETSINF, Universitat Politècnica de València, 16 de gener de 2024

Grup, cognoms i nom: 2,

Problema sobre regressió logística

La següent taula presenta un conjunt de 2 mostres d'entrenament de 2 dimensions procedents de 2 classes:

n	x_{n1}	x_{n2}	c_n
1	1	0	2
2	1	1	1

Addicionalment, la següent taula representa una matriu de pesos inicials amb els pesos de cadascuna de les classes per columnes::

\mathbf{w}_1	\mathbf{w}_2
0.	0.
0.	0.
0.25	-0.25

Es demana:

- (0.5 punts) Calcula el vector de logits associat a cada mostra d'entrenament.
- (0.25 punts) Aplica la funció softmax al vector de logits de cada mostra d'entrenament.
- (0.25 punts) Classifica cadascuna de les mostres d'entrenament. En cas d'empat, tria qualsevol classe.
- (0.5 punts) Calcula el gradient de la funció NLL en el punt de la matriu de pesos inicials.
- (0.5 punts) Actualitza la matriu de pesos inicials aplicant descens per gradient amb factor d'aprenentatge $\eta = 1.0$.

Solució:

- Vector de logits per a cada mostra d'entrenament:

n	a_{n1}	a_{n2}
1	0.	0.
2	0.25	-0.25

- Aplicació de la funció softmax:

n	μ_{n1}	μ_{n2}
1	0.5	0.5
2	0.62	0.38

- Classificació de cada mostra:

n	$\hat{c}(x_n)$
1	2
2	1

- Gradient:

\mathbf{g}_1	\mathbf{g}_2
0.06	-0.06
0.06	-0.06
-0.19	0.19

- Matriu de pesos actualitzada:

\mathbf{w}_1	\mathbf{w}_2
-0.06	0.06
-0.06	0.06
0.44	-0.44