Examen del Bloc 2 de Sistemes Intel·ligents (tipus B)

ETSINF, UPV, 10 de desembre de 2018. Puntuació: encerts - errors/3.

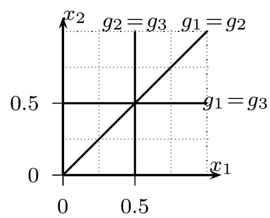
1 ☐ Quina de les següents distribucions de probabilitat *no pot* deduir-se a partir de la prob. conjunta $P(x, y, z)$?:
A) $P(x \mid y)$.
B) $P(z)$.
C) $P(z \mid x, y)$.
D) Tota distribució en la qual intervinga qualsevol combinació d'aquestes variables pot deduir-se de $P(x, y, z)$.

2 ☐ Siga un problema de classificació en quatre classes, $C = \{a, b, c, d\}$, on les quatre classes són equiprobables, i siga y un fet o dada. La decisió òptima de classificació per a y és la classe a amb una probabilitat a posteriori de 0.30. Quina de les següents afirmacions és correcta?
A) La probabilitat d'error és menor que 0.50.
B) $P(Y = y \mid C = a) \; = \; 0.3 \cdot P(Y = y) \; / \; 0.25$.
C) $P(C = a \mid Y = y) \; > \; P(C = b \mid Y = y) + P(C = c \mid Y = y) + P(C = d \mid Y = y)$.
D) Cap de les anteriors.

3 ☐ Suposeu que tenim dues caixes amb 40 galetes cadascuna. La primera caixa conté 10 galetes de xocolate i 30 sense xocolate. La segona caixa conté 20 galetes de cada tipus. Ara suposeu que es tria una caixa a l'atzar, i després una galeta a l'atzar de la caixa triada. Si la galeta triada no és de xocolate, la probabilitat P que procedisca de la primera caixa és:
A) $3/4 \leq P \leq 4/4$.
B) $2/4 \leq P < 3/4$.
C) $1/4 \leq P < 2/4$.
D) $0/4 \leq P < 1/4$.

4 ☐ Siga $\mathbf{x} = (x_1, \ldots, x_D)^t$, $D > 1$, un objecte representat mitjançant un vector de característiques D-dimensional a classificar en una C de classes. Indica quin dels següents classificadors *no* és d'error mínim:
A) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C} \; p(x_1, c)\, p(x_2, \ldots, x_D \mid x_1, c)$
B) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C} \; p(c)\, p(x_1, \ldots, x_D \mid c)$
C) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C} \; p(c \mid x_1)\, p(x_2, \ldots, x_D \mid x_1, c)$
D) $c(\mathbf{x}) = \arg\max_{c=1,\ldots,C} \; p(x_1 \mid c)\, p(x_2, \ldots, x_D \mid x_1, c)$

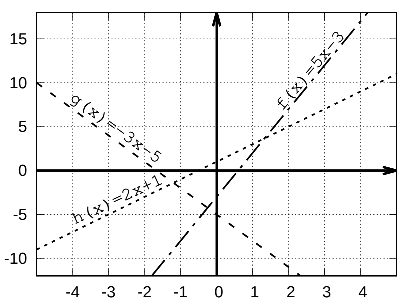
5 ☐ En la figura de la dreta es representen les fronteres de decisió d'un classificador en 3 classes. Quins dels següents vectors de pesos defineixen aquestes fronteres?



A) $\mathbf{w}_1 = (0, 0, 1)^t \quad \mathbf{w}_2 = (0, 1, 0)^t$ i $\mathbf{w}_3 = (0.5, 0, 0)^t$
B) $\mathbf{w}_1 = (0, 0, 1)^t \quad \mathbf{w}_2 = (0, 1, 0)^t$ i $\mathbf{w}_3 = (1, 0, 0)^t$
C) $\mathbf{w}_1 = (0.5, 0, 0)^t \; \mathbf{w}_2 = (0, 1, 0)^t$ i $\mathbf{w}_3 = (0, 0, 1)^t$
D) $\mathbf{w}_1 = (0, 0, 1)^t \quad \mathbf{w}_2 = (1, 0, 0)^t$ i $\mathbf{w}_3 = (0, 1, 0)^t$

6 ☐ Siga un classificador lineal per a dues classes, $\circ$ i $\bullet$, de vectors de pesos $\mathbf{a}_\circ = (2, -5, 4)^t$ i $\mathbf{a}_\bullet = (5, 1, 1)^t$, respectivamente. Quina de les següents afirmacions és correcta?
A) El punt $\mathbf{x}' = (1, 2)^t$ pertany a la classe $\circ$.
B) El punt $\mathbf{x}' = (-2, 0)^t$ es troba a la frontera de decisió.
C) Els vectors de pesos $\mathbf{a}_\circ = (3, 4, 1)^t$ i $\mathbf{a}_\bullet = (2, 2, 2)^t$ defineixen la mateixa frontera de decisió que els de l'enunciat.
D) Els vectors de pesos $\mathbf{a}_\circ = (-2, 5, -4)^t$ i $\mathbf{a}_\bullet = (-5, -1, -1)^t$ defineixen un classificador equivalent al de l'enunciat.

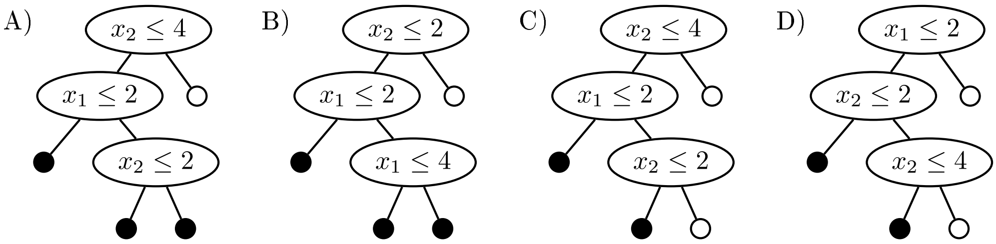
7 ☐ En la figura de la dreta es mostren les funcions discriminants lineals resultants d'entrenar un classificador amb l'algorisme Perceptró i un conjunt de punts de $\mathbb{R}$. Les funcions obtingudes són: $g(x) = -3x - 5$, $h(x) = 2x + 1$ i $f(x) = 5x - 3$. Indica quines són les fronteres de decisió correctes entre $g(x)$ i $h(x)$, i entre $h(x)$ i $f(x)$:

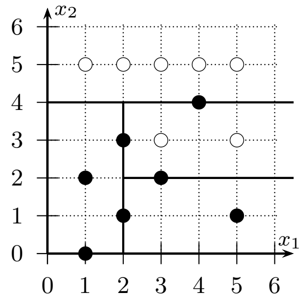


A) $x = -5/3$ i $x = 3/5$.
B) $x = -1/2$ i $x = 3/5$.
C) $x = -5/3$ i $x = 4/3$.
D) $x = -6/5$ i $x = 4/3$.

8 ☐ Indica quina de les següents afirmacions referents a l'algorisme Perceptró (al que direm P) és *certa* quan s'aplica a l'aprenentatge amb una mostra de vectors etiquetats S:
A) El nombre de vectors de S ben classificats amb els pesos obtinguts en cada iteració de P és major que el número vectors ben classificats en la iteració anterior.
B) P sempre convergeix en un nombre finit d'iteracions, encara que és possible que els pesos finalment obtinguts no classifiquen correctament a tots els vectors de S.
C) Com més gran és S, major és el nombre d'iteracions que necessita P per a convergir.
D) Si la mostra d'aprenentatge és linealment separable, P acaba després d'un nombre finit d'iteracions i els pesos resultants permeten classificar S sense errors.

9 ☐ Donat el conjunt de mostres bidimensionals de 2 classes (○ i ●) de la figura de la dreta, quin dels següents arbres de classificació és coherent amb la partició representada?





10 ☐ Siga un problema de classificació en 3 classes (A, B i C) per al qual es disposa de 6 dades d'aprenentatge representades mitjançant vectors de característiques tridimensionals (veure taula a la dreta). Si desitgem aplicar l'algorisme d'aprenentatge d'arbres de classificació amb aquestes dades, quin és el nombre N *de splits diferents* que cal explorar en el node arrel de l'arbre? Nota: no han de tenir-se en compte els *splits* que donen lloc a nodes buits.

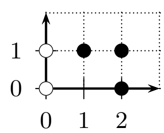
n	1	2	3	4	5	6
x_{n1}	0	1	0	1	0	1
x_{n2}	1	1	2	2	3	3
x_{n3}	0	0	2	3	2	3
c_n	A	A	B	B	C	C

A) $0 \le N < 2$.
B) $2 \le N < 4$.
C) $4 \le N < 6$.
D) $6 \le N$.

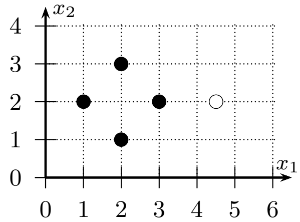
11 ☐ Tenim un problema de classificació en tres classes, $C = \{a, b, c\}$ per a objectes representats en un espai de dues dimensions ($\mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$). Tenim les següents quatre mostres: $\mathbf{y_1} = (4,1)^t$, pertany a la classe a; $\mathbf{y_2} = (1,2)^t$ i $\mathbf{y_3} = (2,3)^t$ pertanyen a la classe b; i $\mathbf{y_4} = (5,1)^t$ pertany a la classe c. Volem construir un arbre de classificació i l'algorisme ha aconseguit un node t que inclou les 4 dades esmentades. Utilitzant la reducció de la impuresa (en termes d'entropia) per a mesurar la qualitat d'un *split,* indica quina de les següents afirmacions és correcta:
A) $\Delta\mathcal{I}(1,2,\,t) > \Delta\mathcal{I}(2,2,\,t)$.
B) $\Delta\mathcal{I}(1,2,\,t) > \Delta\mathcal{I}(2,1,\,t)$.
C) $\Delta\mathcal{I}(2,2,\,t) > \Delta\mathcal{I}(2,1,\,t)$.
D) $\Delta\mathcal{I}(1,4,\,t) = 0$.

12 ☐ Siga T un arbre de classificació construït mitjançant l'algorisme ADC a partir d'una mostra de vectors etiquetats S. Indica quina de les següents afirmacions és *falsa*:
A) Per a tot node t de T, la seua impuresa és igual a la suma de les impureses dels seus nodes fills, t_R i t_L.
B) Si el paràmetre ϵ és prou xicotet, el nombre de vectors de S que T classifica incorrectament pot ser tan xicotet com es vulga.
C) El nombre *de splits* possibles en qualsevol node de T és sempre menor o igual que $D \cdot |S|$, on D és la dimensió dels vectors de S.
D) Encara que T sol ser un arbre aproximadament ben equilibrat, la seua altura pot ser major que $\log_2 |S|$.

13 ☐ En la figura de la dreta es mostra una partició de 5 punts bidimensionals de 2 clústers. La transferència del punt $(1,1)^t$ del clúster ● al clúster ○
A) no altera la SEQ.
B) produeix un increment en la SEQ.
C) produeix un decrement en la SEQ.
D) produeix una SEQ negativa.



14 ☐ En la figura de la dreta es representen 5 mostres bidimensionals particionades inicialment en dos clústers (● i ○). Quin seria el resultat de l'aplicació d'una iteració de l'algorisme C-mitjanes en la seua versió convencional?, i en la versió de Duda i Hart (D&H)?



A) Ambdues versions transfereixen la mostra $(3,2)$.
B) Cap de les dues versions transfereix la mostra $(3,2)$.
C) Només la versió convencional transfereix la mostra $(3,2)$.
D) Només la versió D&H transfereix la mostra $(3,2)$.

15 ☐ S'aplica l'algorisme C-mitjanes de Duda i Hart a un conjunt de N vectors no etiquetats i s'obté una partició de dit conjunt en C subconjunts disjunts de suma d'errors quadràtics, SEQ, igual a R. Quina de les següents afirmacions és *falsa*?:

A) Si $C = N$, $R = 0$.
B) Si $C \ge N/2$, $R = 0$.
C) Si $C \le N$, C-mitjanes acaba en un nombre finit d'iteracions i R és un mínim local de la SEQ.
D) Cap de les anteriors.