

Examen del Bloc 2 de Sistemes Intel·ligents
ETSINF, Universitat Politècnica de València, 13 de gener de 2016

Cognoms:

Nom:

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

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

1 ☒ D Quina de les següents expressions és correcta?

A) $P(x | y) = \frac{1}{P(z)} \sum_x P(x, y, z).$

B) $P(x | y) = \frac{1}{P(z)} \sum_z P(x, y, z).$

C) $P(x | y) = \frac{1}{P(y)} \sum_x P(x, y, z).$

D) $P(x | y) = \frac{1}{P(y)} \sum_z P(x, y, z).$

2 ☒ A Un metge sap que:

- La malaltia de la meningitis causa rigidesa de coll en 70% un dels casos.
- La probabilitat a priori de que un pacient tinga meningitis és 1 / 100 000.
- La probabilitat a priori de que un pacient tinga rigidesa de coll és del 1%.

Amb base en el coneixement anterior, la probabilitat P de que un pacient amb rigidesa de coll tinga meningitis és:

A) $0.000 \leq P < 0.001.$ $P = P(m | r) = \frac{P(m) P(r|m)}{P(r)} = \frac{1/100\,000 \cdot 70/100}{1/100} = 0.0007$

B) $0.001 \leq P < 0.002.$

C) $0.002 \leq P < 0.003.$

D) $0.003 \leq P.$

3 ☒ D Considereu un problema de classificació convencional, açò és, de C classes i objectes representats mitjançant vectors D -dimensionals de característiques reals. En termes generals, podem dir que el problema serà més difícil...

- A) com més menuts siguen C i D .
- B) com més menut siga C i més gran siga D .
- C) com més gran siga C i més menut siga D .
- D) com més grans siguen C i D .

4 ☒ B Es té un problema de classificació per al qual s'han après dos classificadors diferents, c_A i c_B . La probabilitat d'error de c_A s'ha estimat empíricament, a partir d'un cert conjunt de 100 mostres de test, obtenint-se un valor de $\hat{p}_A = 0.10$ (10%). La probabilitat d'error de c_B s'ha estimat anàlogament, si bé en aquest cas s'ha emprat un conjunt de test diferent, compost per 200 mostres, obtenint-se també un 10% d'error ($\hat{p}_B = 0.10$). Amb base en aquestes estimacions, podem afirmar que, per a un nivell de confiança del 95%:

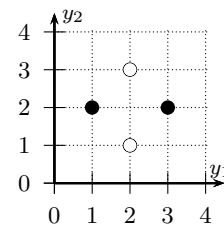
A) Els intervals de confiança de \hat{p}_A i \hat{p}_B seran idèntics.

B) L'interval de confiança de \hat{p}_A serà major que el de \hat{p}_B . $I_A = \hat{p}_A \pm 1.96 \sqrt{\frac{\hat{p}_A (1-\hat{p}_A)}{100}} = 0.10 \pm 0.06$

C) L'interval de confiança de \hat{p}_B serà major que el de \hat{p}_A . $I_B = \hat{p}_B \pm 1.96 \sqrt{\frac{\hat{p}_B (1-\hat{p}_B)}{200}} = 0.10 \pm 0.04$

D) Els intervals de confiança de \hat{p}_A i \hat{p}_B són en aquest cas irrelevants ja que les taxes d'error estimades coincideixen.

- 5 **C** En la figura de la dreta es representen quatre mostres d'aprenentatge bidimensionals de 2 classes: \circ i \bullet . A aquestes mostres se'ls aplica l'algorisme Perceptró amb pesos inicials $\mathbf{a}_\circ = (0, 1, 0)^t$ i $\mathbf{a}_\bullet = (0, 0, 1)^t$, una constant d'aprenentatge $\alpha > 0$ i un marge b . Indica quina de les següents afirmacions és correcta:



- A) L'algorisme convergirà per a algun $b > 0$.
- B) L'algorisme solament pot convergir si $b \leq 0$.
- C) Si $b > 0$, no hi ha convergència, però es pot ajustar el valor d' α tal que, després d'un nombre finit d'iteracions, s'obtinguen bones solucions (amb 25% d'error de resubstitució).
- D) L'algorisme no és aplicable a aquestes mostres perquè no són linealment separables.
- 6 **B** Quin seria el nombre mínim d'errors d'un classificador lineal amb el conjunt de mostres de la qüestió anterior?
- A) 0.
- B) 1.
- C) 2.
- D) 3.

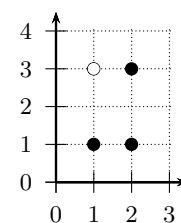
- 7 **B** Donat un classificador lineal de 2 classes \circ i \bullet definit pel seu conjunt de pesos $\mathbf{a}_\circ = (3, 1, 1)^t$ i $\mathbf{a}_\bullet = (1, 2, 1)^t$ (en notació homogènia, la primera component de la qual és el terme independent de la funció lineal corresponent). Quina de les següents afirmacions és correcta?

- A) Com que hi ha dos vectors de pesos i l'espai de representació és bidimensional, tindrem 4 regions de decisió.
- B) Els vectors de pesos $\mathbf{a}_\circ = (2, -2, -2)^t$ i $\mathbf{a}_\bullet = (-2, 0, -2)^t$ determinen la mateixa frontera de decisió que la del classificador donat. L'equació de la frontera és: $\mathbf{a}_\circ^t \mathbf{y} = \mathbf{a}_\bullet^t \mathbf{y}$. En tots dos casos s'obté: $y_1 = 2$.
- C) Un classificador equivalent al donat és el definit per $\mathbf{a}_\circ = (1, 2, 1)^t$ i $\mathbf{a}_\bullet = (3, 1, 1)^t$. Regions de decisió oposades.
- D) Com que els vectors de pesos són de tres dimensions, la frontera ve donada per l'equació d'un plànol en \mathbb{R}^3 .

- 8 **D** Supposeu que estem aplicant l'algorisme d'aprenentatge d'arbres de classificació per a un problema de dues classes, A i B . L'algorisme ha arribat a un node t que inclou dues dades: una de la classe A i una altra de la classe B . La impuresa de t , $\mathcal{I}(t)$, mesurada com l'entropia de la distribució emprírica de les probabilitats a posteriori de les classes a t , és:

- A) $\mathcal{I}(t) < 0.0$.
- B) $0.0 \leq \mathcal{I}(t) < 0.5$.
- C) $0.5 \leq \mathcal{I}(t) < 1.0$.
- D) $1.0 \leq \mathcal{I}(t)$. $\mathcal{I}(t) = -\hat{P}(A | t) \log_2 \hat{P}(A | t) - \hat{P}(B | t) \log_2 \hat{P}(B | t) = -\frac{1}{2} \log_2 \frac{1}{2} - \frac{1}{2} \log_2 \frac{1}{2} = 1$

- 9 **D** La figura a la dreta mostra una partició de 4 punts bidimensionals en 2 clústers (representats mitjançant els símbols \bullet i \circ). La suma d'errors quadràtics (SEQ) d'aquesta partició és $J = \frac{30}{9}$. La transferència del punt $(2, 3)^t$ del clúster \bullet al \circ condueix a un increment de la SEQ, ΔJ , tal que:



- A) $\Delta J > 0$.
- B) $0 \geq \Delta J > -1$.
- C) $-1 \geq \Delta J > -2$.
- D) $-2 \geq \Delta J$. $\Delta J = -\frac{21}{9} = -2.33$ ($J = \frac{30}{9} \rightarrow J = 1$)

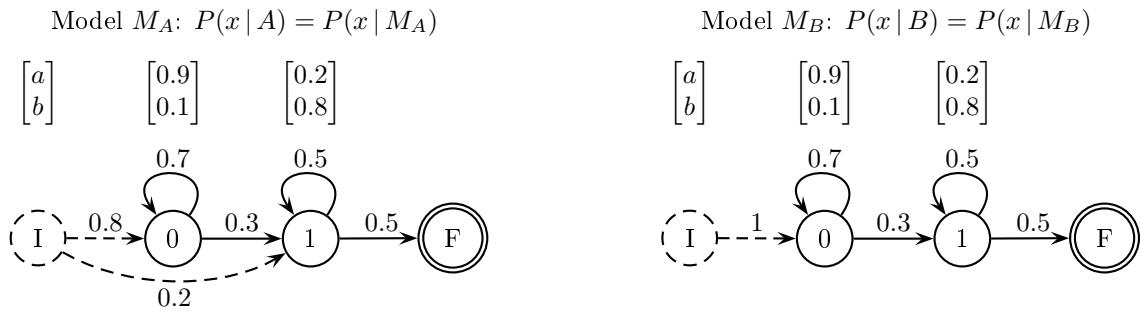
- 10 **B** Dues versions ben conegudes de l'algorisme C -mitjanes són la de *Duda i Hart* (DH) i la "popular". Suposant que ambdues versions s'apliquen a partir d'una mateixa partició inicial, indica quina de les següents afirmacions sobre els seus resultats és certa:

- A) Ambdues versions obtindran la mateixa partició optimitzada.
- B) La versió DH obtindrà una partició final que no podrà millorar-se mitjançant la versió popular.
- C) La versió popular obtindrà una partició final que no podrà millorar-se mitjançant la versió DH.
- D) La partició final obtinguda mitjançant la versió DH podrà millorar-se mitjançant la versió popular, i viceversa.

- 11 **A** Donat el model de Markov M_A de la pregunta 12, l'aproximació de Viterbi a la probabilitat exacta que aquest model assigna a la cadena “bba” és:

- A) 0.003200. $\tilde{P}(bba, q_1 q_2 q_3 = 111 \mid M_A) = 0.2 \cdot 0.8 \cdot 0.5 \cdot 0.8 \cdot 0.5 \cdot 0.2 \cdot 0.5 = 0.0032$
 B) 0.004328.
 C) 0.006400.
 D) Cap dels resultats anteriors és correcte.

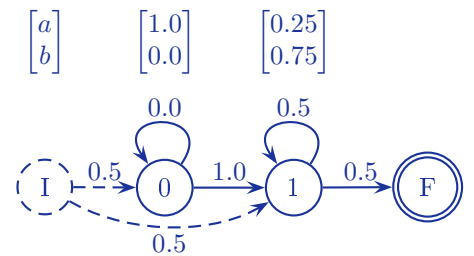
- 12 **B** Es té un problema de classificació en dues classes equiprobables (A i B) d'objectes representats mitjançant cadenes de símbols de l'alfabet $\Sigma = \{a, b\}$. Les funcions de probabilitat condicional de les classes vénen caracteritzades pels models de Markov:



Per mínima probabilitat d'error, la cadena “bba” quedaria classificada en la classe:

- A) Indistintament en A o B ja que les classes són equiprobables.
 B) En la classe A. $\hat{c} = \arg \max_c P(c \mid \text{“bba”}) = \arg \max_c P(c)P(\text{“bba”} \mid c) = \arg \max_c P(\text{“bba”} \mid c)$
 C) En la classe B. $P(\text{“bba”} \mid A) \approx \tilde{P}(\text{“bba”} \mid A) = 0.0032 \gg P(\text{“bba”} \mid B) \approx \tilde{P}(\text{“bba”} \mid B) = 0.0012 \rightarrow \hat{c} = A$
 D) No es pot determinar ja que M_B no compleix les condicions de normalització.
- 13 **C** Donat el model de Markov M_A de la pregunta 12, si apliquem l'algorisme *forward* amb la cadena “bba”, es compleix que:
- A) $\alpha(q = 1, t = 3) = \alpha(q = 0, t = 2) \cdot A_{01} \cdot B_{1a}$.
 B) $\alpha(q = 1, t = 3) = \alpha(q = 1, t = 2) \cdot A_{11} \cdot B_{1a}$.
 C) $\alpha(q = 1, t = 3) = \alpha(q = 0, t = 2) \cdot A_{01} \cdot B_{1a} + \alpha(q = 1, t = 2) \cdot A_{11} \cdot B_{1a}$.
 D) $\alpha(q = 1, t = 3) = \alpha(q = 0, t = 2) \cdot A_{01} \cdot B_{1a} \cdot \alpha(q = 1, t = 2) \cdot A_{11} \cdot B_{1a}$.
- 14 **D** Donat el model de Markov M_A de la pregunta 12, després d'una iteració de re-estimació per Viterbi a partir de les cadenes d'entrenament “bba” i “ab”, es compleix que:

- A) $\pi_0 = 1$.
 B) No es produeix cap canvi en el model.
 C) Totes les probabilitats de transició modifiquen el seu valor.
 D) L'estat 0 té algunes probabilitats d'emissió i/o transició nul·les.



- 15 **B** El model de Markov de conjunt d'estats $Q = \{0, 1, F\}$ i alfabet $\Sigma = \{a, b\}$, estimat mitjançant una inicialització amb una segmentació lineal a partir de les cadenes d'entrenament “bbaa” i “ab”:

- A) Té algunes probabilitats d'emissió nul·les.
 B) Compleix que $A_{00} = A_{11}$ i $A_{01} = A_{1F}$.
 C) Compleix que $\pi_0 = \pi_1$.
 D) Compleix que $B_{0a} = B_{1a}$.

