

Intelligent Systems - Final Exam (Block 2): Test (1.75 points)

ETSINF, Universitat Politècnica de València, January 17th, 2023

Group, surname(s) and name: 2,

Tick only one choice among the given options. Score: $\max(0, (\text{correct_answers} - \text{wrong_answers} / 3) \cdot 1.75 / 9)$.

- 1 ☒ B Suppose we have two boxes containing 40 oranges in the first box and 80 oranges in the second box. The first box contains 26 Navelina oranges and 14 Caracara oranges. The second box contains three times more oranges Navelina than Caracara. Let's assume we randomly pick a box, and then we randomly take one orange from the selected box. If the picked orange is Navelina, which is the probability P that the orange is picked out of the first box? $P = 0.46$

- A) $0/4 \leq P < 1/4$.
 B) $1/4 \leq P < 2/4$.
 C) $2/4 \leq P < 3/4$.
 D) $3/4 \leq P \leq 4/4$.

- 2 ☒ D For a four-class classification problem of objects of type $\mathbf{x} = (x_1, x_2)^t \in \{0, 1\}^2$, we have the probability distributions shown in the table. Show the interval of the Bayes probability of error, ε^* :

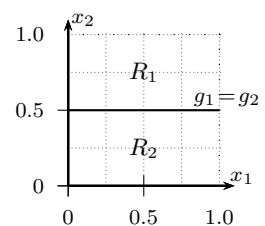
- A) $\varepsilon^* < 0.40$.
 B) $0.40 \leq \varepsilon^* < 0.45$.
 C) $0.45 \leq \varepsilon^* < 0.50$.
 D) $0.50 \leq \varepsilon^*$.

\mathbf{x}		$P(c \mathbf{x})$				$P(\mathbf{x})$
x_1	x_2	$c=1$	$c=2$	$c=3$	$c=4$	
0	0	0.2	0.3	0.4	0.1	0.1
0	1	0.3	0.4	0.2	0.1	0.3
1	0	0.3	0.3	0.1	0.3	0.1
1	1	0.3	0.4	0.1	0.2	0.5

$$\varepsilon^* = 0.61$$

- 3 ☒ A The figure on the right represents the decision boundary and the two regions of a binary classifier. Which of the following weight vectors (in homogeneous notation) defines a classifier equivalent to the one of the figure?

- A) $\mathbf{w}_1 = (0, 0, 2)^t$ and $\mathbf{w}_2 = (1, 0, 0)^t$.
 B) $\mathbf{w}_1 = (0, 0, -2)^t$ and $\mathbf{w}_2 = (-1, 0, 0)^t$.
 C) $\mathbf{w}_1 = (1, 0, 0)^t$ and $\mathbf{w}_2 = (0, 0, 2)^t$.
 D) All the above weight vectors define an equivalent classifier.



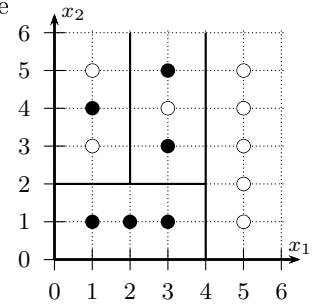
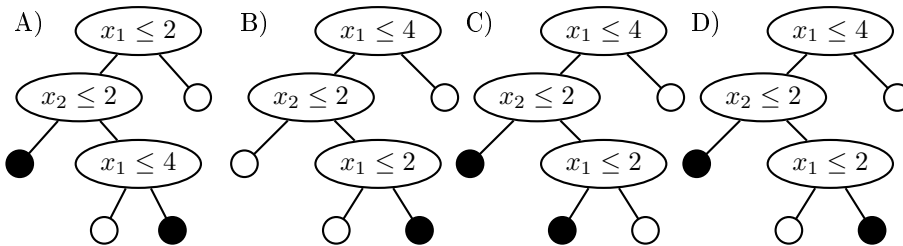
4 **B** Let's suppose that we are applying the Perceptron algorithm, with learning rate $\alpha = 1$ and margin $b = 0.1$, to a set of 4 bidimensional learning samples for a problem of 4 classes, $c = 1, 2, 3, 4$. At a given moment in the execution of the algorithm, we have obtained the weight vectors $\mathbf{w}_1 = (-2, -8, -12)^t$, $\mathbf{w}_2 = (-2, -6, -4)^t$, $\mathbf{w}_3 = (-2, -6, -4)^t$, $\mathbf{w}_4 = (-2, -8, -8)^t$. Assuming that the sample $(\mathbf{x}, c) = ((5, 3)^t, 2)$ is then going to be processed, how many weight vectors will be modified?

- A) 0
- B) 2
- C) 3
- D) 4

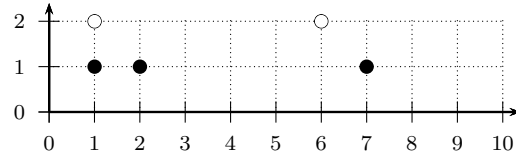
5 **C** Suppose we apply the classification tree algorithm for a two-class problem $c = A, B$. The algorithm reaches a node t whose impurity, measured as the entropy impurity of the posterior probability of the classes in the node t , is $I = 0.72$. Which is the number of samples in each class at node t ?

- A) 8 in class A and 64 in class B
- B) 16 in class A and 128 in class B
- C) 16 in class A and 64 in class B
- D) 8 in class A and 128 in class B

6 **D** Given the two-class (\circ and \bullet) samples of the figure on the right, which of the following classification trees is coherent with the partition of the figure?



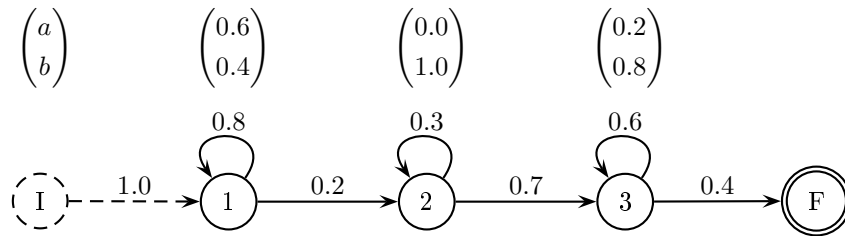
- 7 A The figure below shows a partition of 5 two-dimensional points in 2 clusters, \bullet and \circ :



If point $(7, 1)^t$ is transferred from one cluster to the other, a variation of the Sum of Square Errors (SSE) is produced, $\Delta J = J - J'$ (SSE after the transfer minus SSE before the transfer), such that:

- A) $\Delta J < -7$. $\Delta J = 21.8 - 33.2 = -11.3$
 B) $-7 \leq \Delta J < 0$.
 C) $0 \leq \Delta J < 7$.
 D) $\Delta J \geq 7$.

- 8 A The figure below shows a graphical representation of a Markov model M :



How many different strings of length 2 starting with the symbol a can be generated with M ? 0

- A) None.
 B) One.
 C) Two.
 D) More than two.

- 9 D Let M be a Markov model with states $Q = \{1, 2, F\}$; alphabet $\Sigma = \{a, b\}$; initial probabilities, $\pi_1 = \frac{1}{2}, \pi_2 = \frac{1}{2}$; matrix A for transition probabilities, matrix B for emission probabilities, and Forward matrix α :

A	1	2	F
1	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{3}{7}$
2	$\frac{1}{7}$	$\frac{3}{7}$	$\frac{3}{7}$

B	a	b
1	$\frac{3}{6}$	$\frac{3}{6}$
2	$\frac{1}{4}$	$\frac{3}{4}$

α	b	b
1	$\frac{1}{4}$	α_{12}
2	$\frac{3}{8}$	α_{22}

Which are the corresponding values for α_{12} and α_{22} ? $\alpha_{12} = \frac{1}{4} \cdot \frac{2}{7} \cdot \frac{3}{6} + \frac{3}{8} \cdot \frac{1}{7} \cdot \frac{3}{6}$, $\alpha_{22} = \frac{1}{4} \cdot \frac{2}{7} \cdot \frac{3}{4} + \frac{3}{8} \cdot \frac{3}{7} \cdot \frac{3}{4}$

- A) $\alpha_{12} = \frac{1}{28}, \alpha_{22} = \frac{27}{224}$
 B) $\alpha_{12} = \frac{1}{16}, \alpha_{22} = \frac{27}{224}$
 C) $\alpha_{12} = \frac{1}{28}, \alpha_{22} = \frac{39}{224}$
 D) $\alpha_{12} = \frac{1}{16}, \alpha_{22} = \frac{39}{224}$

Intelligent Systems - Final Exam (Block 2): Problem (2 points)

ETSINF, Universitat Politècnica de València, January 17th, 2023

Group, surname(s) and name: 2,

Problem: Viterbi algorithm

Let M be a Markov model with states $Q = \{1, 2, F\}$; alphabet $\Sigma = \{a, b\}$; prior probabilities $\pi_1 = \frac{1}{2}, \pi_2 = \frac{1}{2}$; and transition and emission probabilities:

A	1	2	F
1	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$
2	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{2}{8}$

B	a	b
1	$\frac{2}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	$\frac{2}{3}$

Answer the following questions:

- (1 point) Show a trace of the *Viterbi* algorithm to obtain the most probable sequence of states for generating the string **aa** with M .
- (1 point) Given the training pairs, string - Viterbi sequence, (**ab**, 12F) and (**bbab**, 2212F) together with the string **aa** and its Viterbi sequence computed in the previous question, apply one iteration of the Viterbi re-estimation algorithm to re-estimate the parameters of M .

Solution:

- Viterbi trace for the string **aa** (states 1 and 2 are represented as 0 and 1, respectively):

```

      a      a
0 0.333389 0.041680 0.020842
1 0.166692 0.037048 0.009263
Q:      1      0

```

- Viterbi re-estimation from the pair **aa** and 21F as computed in the previous question, together with the given pairs (**ab**, 12F) and (**bbab**, 2212F), we obtain the re-estimated parameters:

π	1	2
	$\frac{1}{3}$	$\frac{2}{3}$

A	1	2	F
1	$\frac{0}{3}$	$\frac{2}{3}$	$\frac{1}{3}$
2	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

B	a	b
1	$\frac{3}{3}$	$\frac{0}{3}$
2	$\frac{1}{5}$	$\frac{4}{5}$

Through the application of one more Viterbi re-estimation iteration, it is easy to see that the algorithm converges to the former model.