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

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

Group, surname(s) and name: 1,

Tick only one choice among the given options. Score: max(0, (correct_answers − wrong_answers / 3) · 1.75 / 9).

1 $\boxed{\text{A}}$ Suppose we have two boxes containing 40 oranges in the first box and 80 oranges in the second box. The first box contains 9 Navelina oranges and 31 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? $P = 0.23$

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

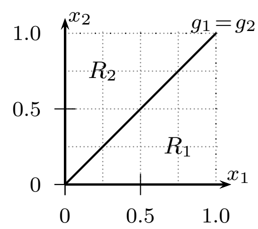
2 $\boxed{\text{D}}$ For a four-class classification problem of objects of type $\mathbf{x} = (x_1, x_2)^t \in \{0,1\}^2$, we have the probability distributions shown in the table. Show the interval of the Bayes probability of error, ε^*:

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

$\mathbf{x}$		$P(c \mid \mathbf{x})$				
x_1	x_2	$c=1$	$c=2$	$c=3$	$c=4$	$P(\mathbf{x})$
0	0	0.1	0.3	0.1	0.5	0
0	1	0.2	0.5	0.3	0	0.1
1	0	0.2	0.4	0.1	0.3	0.3
1	1	0.1	0.3	0.3	0.3	0.6

$\varepsilon^* = 0.65$

3 $\boxed{\text{B}}$ 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) $\mathbf{w}_1 = (0, -2, 0)^t$ and $\mathbf{w}_2 = (0, 0, -2)^t$.
B) $\mathbf{w}_1 = (0, 2, 0)^t$ and $\mathbf{w}_2 = (0, 0, 2)^t$.
C) $\mathbf{w}_1 = (0, 0, 2)^t$ and $\mathbf{w}_2 = (0, 2, 0)^t$.
D) All the above weight vectors define an equivalent classifier.

4 $\boxed{\text{D}}$ 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 $\mathbf{w}_1 = (-2, -2, -6)^t$, $\mathbf{w}_2 = (-2, -2, -6)^t$, $\mathbf{w}_3 = (-2, -4, -4)^t$, $\mathbf{w}_4 = (-2, -4, -4)^t$. Assuming that the sample $(\mathbf{x}, c) = ((4, 5)^t, 2)$ is then going to be processed, how many weight vectors will be modified?

A) 0

B) 2

C) 3

D) 4

5 $\boxed{\text{D}}$ 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?

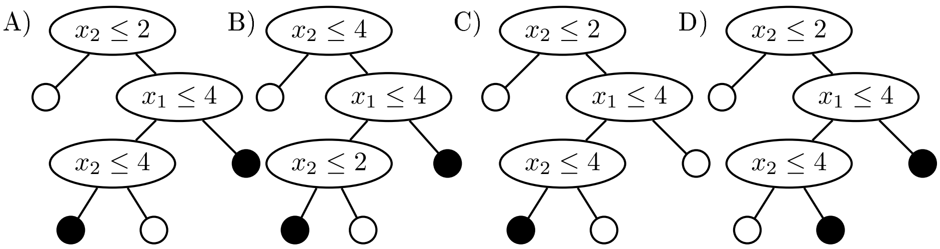
A) 2 in class A and 32 in class B

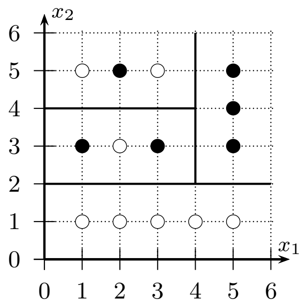
B) 2 in class A and 16 in class B

C) 4 in class A and 32 in class B

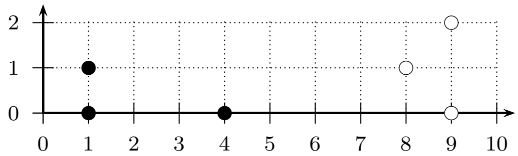
D) 4 in class A and 16 in class B

6 $\boxed{\text{A}}$ 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 $\boxed{D}$ The figure below shows a partition of 6 two-dimensional points in 2 clusters, $\bullet$ and $\circ$:



If point $(1, 0)^t$ 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:

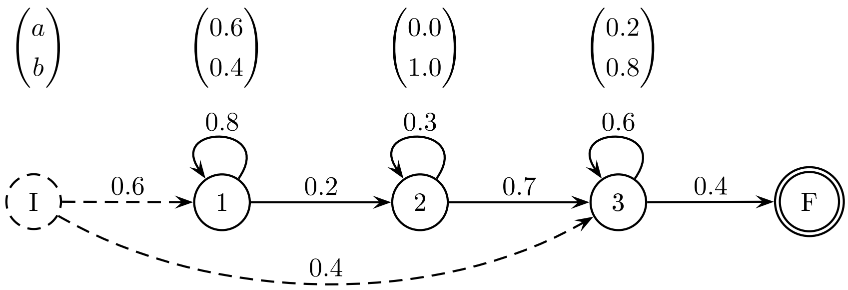
A) $\Delta J < -7$. $\quad\Delta J = 52.5 - 9.3 = 43.2$

B) $-7 \leq \Delta J < 0$.

C) $0 \leq \Delta J < 7$.

D) $\Delta J \geq 7$.

8 $\boxed{D}$ The figure below shows a graphical representation of a Markov model M:



How many different strings of length 3 starting with the symbol a can be generated with M? 4

A) None.

B) One.

C) Two.

D) More than two.

9 $\boxed{C}$ Let M be a Markov model with states $Q = \{1, 2, F\}$; alphabet $\Sigma = \{a, b\}$; initial probabilities, $\pi_1 = \frac{2}{3}, \pi_2 = \frac{1}{3}$; matrix A for transition probabilities, matrix B for emission probabilities, and Forward matrix α:

A	1	2	F
1	$\frac{3}{7}$	$\frac{3}{7}$	$\frac{1}{7}$
2	$\frac{2}{6}$	$\frac{2}{6}$	$\frac{2}{6}$

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

α	b	b
1	$\frac{1}{3}$	α_{12}
2	$\frac{1}{6}$	α_{22}

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

A) $\alpha_{12} = \frac{25}{252}$, $\alpha_{22} = \frac{1}{14}$

B) $\alpha_{12} = \frac{1}{14}$, $\alpha_{22} = \frac{25}{252}$

C) $\alpha_{12} = \frac{25}{252}$, $\alpha_{22} = \frac{25}{252}$

D) $\alpha_{12} = \frac{1}{14}$, $\alpha_{22} = \frac{1}{14}$

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

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

Group, surname(s) and name: 1,

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	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

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

Answer the following questions:

1. (1 point) Show a trace of the *Viterbi* algorithm to obtain the most probable sequence of states for generating the string `ab` with M.
2. (1 point) Given the training pairs, string - Viterbi sequence, (`ba`, `22F`) and (`baa`, `111F`) together with the string `ab` 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:

1. Viterbi trace for the string `ab` (states 1 and 2 are represented as 0 and 1, respectively):

```
            a        b
 0 0.333389 0.055573 0.013894
 1 0.166692 0.055573 0.027789
Q:          0        1
```

2. Viterbi re-estimation from the pair `ab` and `12F` as computed in the previous question, together with the given pairs (`ba`, `22F`) and (`baa`, `111F`), we obtain the re-estimated parameters:

π	1	2
	$\frac{2}{3}$	$\frac{1}{3}$

A	1	2	F
1	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{0}{3}$	$\frac{1}{3}$	$\frac{2}{3}$

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

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