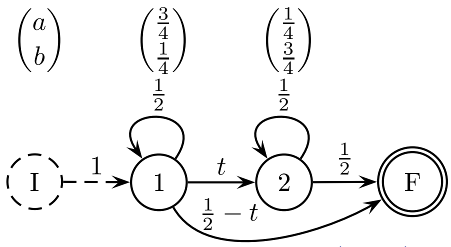
Intelligent Systems - Exam (Block 2): Test (1.75 points)

ETSINF, Universitat Politècnica de València, January 13th, 2022

Group, surname(s) and name: 2,

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

1 $\boxed{\text{B}}$ Let M be the Markov model represented at the right, where t, $0 < t < \frac{1}{4}$, is the transition probability from state 1 to 2. Given the string $x = \texttt{abb}$, the probability for x to be generated through the path $122F$, $P(\texttt{abb}, 122F)$, depends on t. Analogously, the probability for x to be generated through the path $111F$, $P(\texttt{abb}, 111F)$, also depends on t (through the transition probability from state 1 to F). Show in which case $P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F)$:



$$P(\texttt{abb}, 111F) = 1\frac{3}{4}\frac{1}{2}\frac{1}{4}\frac{1}{2}\frac{1}{4}\left(\frac{1}{2} - t\right)$$

$$P(\texttt{abb}, 122F) = 1\frac{3}{4}t\frac{3}{4}\frac{1}{2}\frac{3}{4}\frac{1}{2}$$

$$P(\texttt{abb}, 111F) > P(\texttt{abb}, 122F) \rightarrow t < \frac{1}{20}$$

A) Never.

B) If and only if $0 < t < \frac{1}{20}$.

C) If and only if $0 < t < \frac{1}{10}$.

D) Always, that is, $0 < t < \frac{1}{4}$.

2 $\boxed{\text{D}}$ Given the following 3 nodes of a classification tree with samples belonging to 3 classes:

c	n_1	n_2	n_3
1	2	4	5
2	3	5	3
3	1	1	5

where each row is the number of samples of each class in the node. Which of the following inequalities is true?

A) $\mathcal{I}(n_2) < \mathcal{I}(n_3) < \mathcal{I}(n_1)$

B) $\mathcal{I}(n_1) < \mathcal{I}(n_3) < \mathcal{I}(n_2)$

C) $\mathcal{I}(n_3) < \mathcal{I}(n_2) < \mathcal{I}(n_1)$

D) $\mathcal{I}(n_2) < \mathcal{I}(n_1) < \mathcal{I}(n_3)$

3 $\boxed{\text{B}}$ Given the following dataset to train a classification tree with 5 bidimensional samples that belong to 2 classes:

n	1	2	3	4	5
x_{n1}	3	2	3	2	1
x_{n2}	2	3	5	1	1
c_n	2	2	2	2	1

How many different partitions can be generated at the root node? Do not consider those partitions in which all data samples are assigned to the same child node.

A) 4

B) 5

C) 3

D) 7

4 $\boxed{\text{D}}$ In a problem of probabilistic reasoning corresponding to road trips, with the random variables: Climatology (C):{clear (CLE), cloudy (CLO), rainy (RAI)}; Luminosity (L):{day (DAY), night(NIG)}; Security (S):{secure (SEC), accident (ACC)}. The joint probability of the three random variables is provided by the table:

	DAY			NIG		
$P(s,l,c)$	CLE	CLO	RAI	CLE	CLO	RAI
SEC	0.28	0.21	0.04	0.15	0.09	0.09
ACC	0.02	0.02	0.03	0.02	0.02	0.03

The conditional probability $P(S = \text{SEC} \mid L = \text{DAY}, C = \text{CLO})$ is:

A) 0.230

B) 0.210

C) 0.860

D) 0.913

5 $\boxed{\text{B}}$ Let M be a Markov model with set of states $Q = \{1, 2, F\}$ and alphabet $\Sigma = \{a, b\}$. During the application of an iteration of the Viterbi's reestimation algorithm, a pair "(string, most probable path)" has been obtained for each training string. Then, from all the pairs obtained, the counts (absolute frequencies) of state transitions have been also obtained and are shown at the right table. The *correct* normalization of these counts will result in the table of state transition probabilities:

A	1	2	F
1	1	1	4
2	4	3	1

A)

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

B)

A	1	2	F
1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{4}{6}$
2	$\frac{4}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

C)

A	1	2	F
1	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{4}{5}$
2	$\frac{4}{5}$	$\frac{3}{4}$	$\frac{1}{5}$

D)

A	1	2	F
1	$\frac{1}{14}$	$\frac{1}{14}$	$\frac{4}{14}$
2	$\frac{4}{14}$	$\frac{3}{14}$	$\frac{1}{14}$

6 $\boxed{\text{D}}$ The estimated probability of error of a classifier is 20%. Which is the minimum number of testing samples, M, so that the 95% confidence interval of this estimated probability of error is not higher than $\pm 1\%$; that is, $I = [19\%, 21\%]$: $M = 6147$

A) $M < 2000$.

B) $2000 \leq M < 3500$.

C) $3500 \leq M < 5000$.

D) $M \geq 5000$.

7 $\boxed{\text{A}}$ Given a classifier for 3 classes defined by their weight vectors $\mathbf{w}_1 = (-1, 1, -3)^t$, $\mathbf{w}_2 = (-3, 1, -3)^t$, $\mathbf{w}_3 = (0, -3, -2)^t$ in homogeneous notation, which of the following weight vectors do **not** define a classifier equivalent to the one given?

A) $\mathbf{w}_1 = (1, -1, 3)^t$, $\mathbf{w}_2 = (3, -1, 3)^t$, $\mathbf{w}_3 = (0, 3, 2)^t$
B) $\mathbf{w}_1 = (0, 2, -6)^t$, $\mathbf{w}_2 = (-4, 2, -6)^t$, $\mathbf{w}_3 = (2, -6, -4)^t$
C) $\mathbf{w}_1 = (-2, 2, -6)^t$, $\mathbf{w}_2 = (-6, 2, -6)^t$, $\mathbf{w}_3 = (0, -6, -4)^t$
D) $\mathbf{w}_1 = (1, 1, -3)^t$, $\mathbf{w}_2 = (-1, 1, -3)^t$, $\mathbf{w}_3 = (2, -3, -2)^t$

8 $\boxed{\text{B}}$ Let's suppose that we are applying the Perceptron algorithm, with learning rate $\alpha = 1$ and margin $b = 0.1$, to a set of 3 bidimensional learning samples for a problem of 2 classes. After processing the first 2 samples, the weight vectors $\mathbf{w}_1 = (0, 0, 3)^t$, $\mathbf{w}_2 = (0, 0, -3)^t$ were obtained. Next, the last sample $(\mathbf{x}_3, c_3)$ is processed and the following weight vectors $\mathbf{w}_1 = (-1, -5, -1)^t$, $\mathbf{w}_2 = (1, 5, 1)^t$ are obtained, which of the following samples is that last sample?

A) $((1, 1)^t, 2)$
B) $((5, 4)^t, 2)$
C) $((3, 3)^t, 2)$
D) $((3, 2)^t, 2)$

9 $\boxed{\text{A}}$ We have a partition of a set of 3-dimensional data points into a given number of clusters, $C \geq 2$. Consider the transfer of the data point $\mathbf{x} = (9, 2, 9)^t$ from a cluster i to another one j, $j \neq i$. We know that cluster i contains 4 data points (including $\mathbf{x}$) and cluster j 3. We also know that the centroid (mean) of cluster i is $\mathbf{m}_i = (2, 9, 2)^t$, while that of cluster j is $\mathbf{m}_j = (6, 2, 3)^t$. If the transfer is carried out, an increase of the sum of square errors, ΔJ, will be produced such that: $\Delta J = -162.2$

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

Intelligent Systems - Exam (Block 2): Problems (2 points)

ETSINF, Universitat Politècnica de València, January 13th, 2022

Group, surname(s) and name: 2,

Problem about Forward and Viterbi

Let M be a Markov model with states $Q = \{1, 2, F\}$; alphabet $\Sigma = \{a, b\}$; initial probabilities $\pi_1 = \frac{1}{3}, \pi_2 = \frac{2}{3}$; and transition and emission probabilities:

A	1	2	F
1	$\frac{2}{4}$	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

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

Let x=`ba`. Answer the following questions:

1. $(0,75$ points) Show a trace of the *Forward* algorithm to obtain the probability for M to generate the string x, $P_M(x)$.
2. $(0,75$ points) Show a trace of the *Viterbi* algorithm to obtain the Viterbi aproximation to the probability for M to generate the string x, $\tilde{P}_M(x)$.
3. $(0,25$ points) From the trace shown in the previous question, find a most probable path through which M generates the string x.
4. $(0,25$ points) Find the probability by which M generates the string x through a path different from that found in the previous question.

Solution:

1. *Forward:* $P_M(x) = 39/800 = 0.04875$

	b	a	
1	1/6	13/120	
2	1/3	13/240	
F			39/800

2. Viterbi: $\tilde{P}_M(x) = 1/60 = 0.01667$

	b	a	
1	1/6	1/15	
2	1/3	1/30	
F			1/60

3. Most probable path: $21F$
4. $P_M(x) - \tilde{P}_M(x) = 77/2400 = 0.03208.$