

Machine Learning – Exam - December 16, 2019

Time limit: **2 hours**.

Last Name

First Name

Matricola

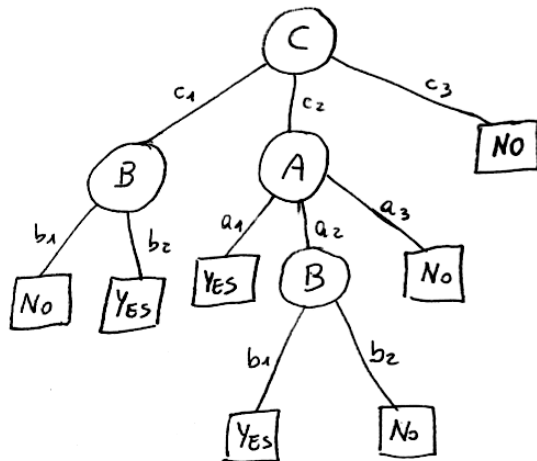
.....

Note: if you are not doing the regular exam for ML 2019/20, write below name of exam, CFU, and academic year (when you were supposed to attend the course). Please specify also if you are an Erasmus student.

.....

EXERCISE 1

Given a classification problem for the function $f : A \times B \times C \rightarrow \{+, -\}$, with $A = \{a_1, a_2, a_3\}$, $B = \{b_1, b_2\}$, $C = \{c_1, c_2, c_3\}$ and the following decision tree T that is the result of a learning algorithm on a given data set:



1. Provide a rule based representation of the tree T .
2. Determine if the tree T is consistent with the following set of samples $S \equiv \{s_1 = \langle a_1, b_1, c_1, No \rangle, s_2 = \langle a_2, b_1, c_2, Yes \rangle, s_3 = \langle a_1, b_2, c_3, Yes \rangle, s_4 = \langle a_2, b_2, c_2, Yes \rangle\}$. Show all the passages needed to get to the answer.

EXERCISE 2

In Bayesian Learning, given a data set D and a hypothesis h , we can express the following relationship between the probability distributions (Bayes theorem):

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)}$$

In this context:

1. define *Maximum a posteriori* (MAP) hypotheses and *Maximum likelihood* (ML) hypotheses.
2. define the concept of *Bayes Optimal Classifier*
3. discuss about practical applicability of the *Bayes Optimal Classifier*

EXERCISE 3

1. Describe the perceptron model for classification and its training rule.
2. Draw a graphical representation of a 2D data set for binary classification and provide a qualitative graphical example of a possible evolution of perceptron training (4 images showing a possible temporal evolution of the solution of the algorithm on the sketched data set, with the last image showing a possible final solution).

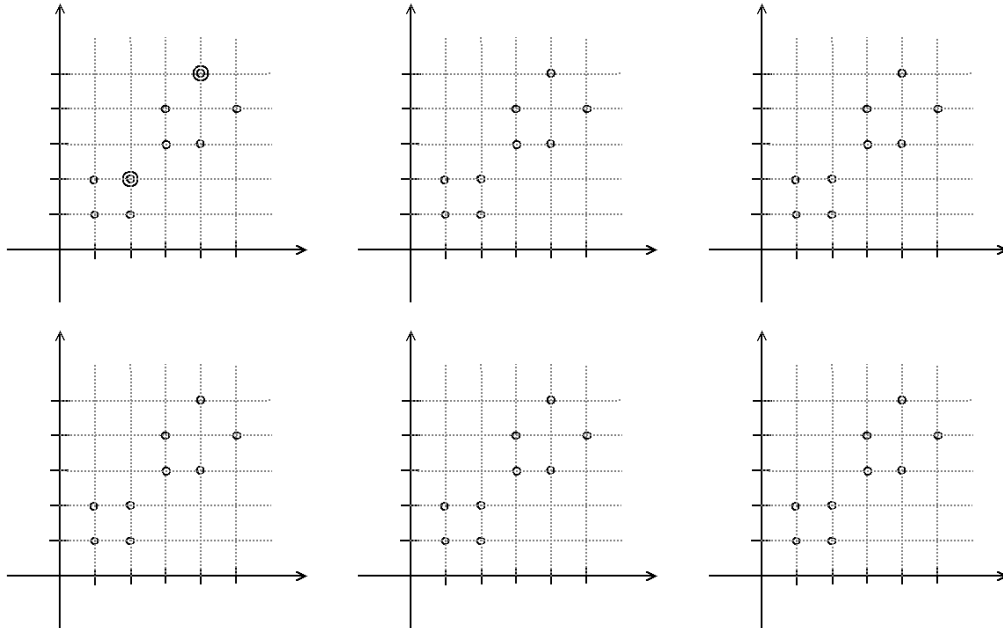
EXERCISE 4

Consider a two-layers ANN which receives in input vectors \mathbf{x} of dimension 128 and produces output vectors \mathbf{y} of dimension 10. The hidden layer of the ANN is composed of 50 units which use the ReLU activation function. The output units use a linear activation function. The weight matrices of the hidden and output layers are denoted W_1 and W_2 , respectively.

1. Provide the dimensions of the weight matrices W_1 and W_2
2. Provide the formula explicitly stating how the values of \mathbf{y} are computed given an input vector \mathbf{x} in terms of the weight matrices and the activation functions (you can ignore the bias terms).

EXERCISE 5

Simulate the execution of K-means in this 2-D data set with $k=2$ and initial centroids indicated by double circles: use one diagram for each step of the algorithm. Describe explicitly how each step is obtained and what is the termination condition of the algorithm. Drawing only the steps is not sufficient.



EXERCISE 6

1. Briefly describe what is the architecture of an autoencoder and its purpose.
2. Draw an example of autoencoder.