Machine Learning – April 12, 2019

Time limit: 2 hours.

Last Name ..... First Name ..... Matricola .....

Note: if you are not doing the exam for ML 2018/19, 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

1. Provide a formal (domain-independent and solution-independent) definition of overfitting.
2. Discuss the problem of overfitting in learning with Decision Trees and illustrate possible solutions to it.

EXERCISE 2

1. Describe the *Naive Bayes Classifier* and highlight the approximation made with respect to the Bayes Optimal Classifier.
2. Provide design and implementation choices for solving the following problem through *Naive Bayes Classifier*:

   Classification of scientific papers in categories according to their main subject. The categories to be considered are: ML (Machine Learning), KR (Knowledge Representation), PL (Planning). Data available for each scientific paper are: title, authors, abstract and publication site (name of the journal and/or of the conference).

EXERCISE 3

Consider a dataset $\mathcal{D} = \{\langle (a_1, s_1), p_1 \rangle, \ldots, \langle (a_N, s_N), p_N \rangle\}$ containing the number of hours a_i a student has attended a course, the number of hours s_i s/he has studied for the course and whether or not s/he has passed the exam $p_i = \{0, 1\}$.

1. Define a model based on **logistic regression** that, given the values of a and s, estimates whether a student passes the exam or not.
2. Discuss which are the parameters of the model that have to be learned based on the given data.
3. What is a suitable error function for learning the parameters of the model?

EXERCISE 4

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

EXERCISE 5

1. Describe the concept of full observability in models representing dynamic systems.
2. Describe the difference between a Markov Decision Process (MDP) and a Partially Observable Markov Decision Process (POMDP), referring to their formal models.
3. Draw and explain the graphical models of MDP and POMDP.

EXERCISE 6

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).