**Machine learning**

**Requirements and questions**

## I. Requirements for project presentation

1. Present the input images used. Specify which image (or images) is linearly separable and which is not.
2. Specify the classifiers (methods) used for solving the assessment task. How many classifiers did you use? (You need to have minimum 3 different classifiers)
3. Show in code the sequence of instructions for creating the classifier (object).
4. Specify the methods used for training (i.e. without splitting, with splitting of data into training and validation datasets, using cross validation) and show in code the sequence of instructions used in any case. You can use 2 or more training methods with the same classifier and make comparisons, or use them with different classifiers)
5. Specify the methods used for performance evaluation (accuracy computed on training and/or testing sets, cross-validation accuracy computed on training and/or testing sets) and be able to present the code which performs these evaluations.
6. Show in code the instruction or the sequence of instructions for prediction.
7. Specify and show in code how did you chose the parameters of your classifiers (default, tune by hand, using grid search cross validation, other methods).
8. Make a report with all your results (table with results and comments about the strengths and drawbacks that you observed to the methods that you have used).
9. Make comparisons, interpret your results and present what you consider more interesting in your report (in your results).

List of possible questions

1. Define the supervised learning problem (of a function). Provide examples of supervised learning problems.
2. Define the unsupervised learning problem (of a function). Provide examples of supervised learning problems.
3. Define the classification problem. Is the classification problem a supervised learning problem of a function? Justify your answer.
4. Define a set of data linearly separable and non-linearly separable.
5. Define the supervised learning problem (of a function). Provide examples of supervised learning problems.
6. Define the unsupervised learning problem (of a function). Provide examples of supervised learning problems.
7. Define the classification problem. Is the classification problem a supervised learning problem of a function? Justify your answer.
8. Explain the following notions used in supervised learning: instance, attribute (feature), target function, hypothesis, training set, noise, bias. For practical problems characterize and identify these elements.
9. Define the noise in supervised learning, enumerate and define the types of noises.
10. Define the bias in supervised learning, enumerate and define the types of bias. Provide an example for each type of bias.
11. Define the inductive learning.
12. How can you evaluate the performances of a supervised learning algorithm (machine learning method)?

13. Define the accurracy of a supervised learning algorithm.
14. Define the precision and recall of a supervised learning algorithm.
15. Define a linear set of data. Give examples.
16. Define a non linear set of data. Give examples.
17. What is cross-validation and what can it be used for?
18. How can you validate a model built using supervised machine learning techniques?
19. What is the overfitting in a supervised learning model?
20. Give general methods for avoiding overfitting.
21. How can you optimize the choice of parameters of a model obtained using a machine learning technique (e.g. Artificial Neural Networks, SVM).
22. What is the INPUT and OUTPUT in a concept learning problem. Define a hypothesis consistent with a training data set D, in a supervised learning problem (in particular in concept learning).
23. What are a positive and a negative instance in a concept learning problem?
24. What is the Version Space in a concept learning problem?
25. How does the Find S algorithm work?
26. How does the Candidate Elimination algorithm work?
27. Which is the representation of a decision tree?
28. Give some appropriate problems for decision tree learning. Justify your choice.
29. How can you decide which attribute will be situated in the root of a decision tree?
30. Present the idea from ID3 algorithm.
31. Which is the inductive bias of ID3 algorithm?
32. How can you validate a decision tree?
33. How does a perceptron work?
34. Give examples of activation functions for a perceptron.
35. Briefly described the Backpropagation algorithm.
36. What does a neural network learn?
37. Give examples of appropriate problems to be solved with neural networks. Specify the input and output.
38. What are the main parameters of a neural network? What are the methods for choosing these parameters (how can you optimize the parameters' choice)?
39. State the version of Bayes' theorem used in Machine Learning (in Baysian learning).
40. What is a maximum a posteriori hypothesis in Baysian learning?
41. How does the Bayes optimal classifier classify an unknown instance?
42. Give the expression of a linear classifier.
43. What is the SVM principle for binary classification of linearly separable data?
44. What are the support vectors in the SVM approach?
45. How does the SVM work in the case of non-linearly separable data?
46. What do you understand by "kernel trick" in the SVM approach for non-linearly separable data?