

MAC0460/5832 – Introduction to Machine Learning

DCC / IME-USP — First semester/2021

Lista 3 – Submission deadline: 20/07/2021

The first part of this list aims to evaluate the level of comprehension about some topics. The second part corresponds to a self-evaluation.

PART I

1. We discussed linear regression and two ways to solve this problem. Explain how a polynomial regression problem could be solved.
2. Considering the context of binary classification problems, discuss similarities and differences between the logistic regression and SVM algorithms.
3. In your opinion, when we confront neural networks and SVM, what are the interesting characteristics of each of them ? Explain.
4. In the context of *machine learning* and according to what we have discussed, what is the difference between validation and test? Explain.
5. What do you understand about *overfitting*? How can we detect *overfitting* and how can we combat it ?

PART II

During the semester, several topics have been discussed. The main topics are listed below:

- Supervised learning × Non-supervised learning
- Discriminative approaches × Generative approaches
- Regression × Classification
- Optimization of a cost function
- Linear regression, its analytic solution, solution based on gradient descent
- Logistic regression, maximization of the likelihood function
- Hoeffding inequality, dichotomies, VC bound
- Bias-variance theory
- Perceptron networks and multilayer neural networks

- Backpropagation algorithm
- SVM, linearly separable case and non-linearly separable case
- Quadratic optimization problems
- Overfitting
- Regularization
- Validation, Cross-validation, Model selection
- Classifier evaluation
- Clustering, dimensionality reduction
- Images, Viola-Jones, Bag-of-words
- Convolutional neural networks
- Others

6. Taking the above list into consideration, evaluate your degree of learning along the semester, by answering the following questions.

- Percentage of topics you consider you have understood sufficiently well [0% a 100%]
- Which is the topic on which you most advanced your knowledge? Briefly discuss it.
- Which are the two topics on which you least advanced your knowledge? Explain why.
- Overall, considering the answers to the previous items, which grade between 0 and 10 would you attribute to your learning?

7. Taking the above list into consideration, evaluate your dedication along the semester, by answering the following questions.

- Percentage of topics you have studied (either attending online classes or watching them of-line, or watching the Caltech videos, or reading the textbook, or studying extra material) [0% a 100%]
- How many QT, Lists, or EP you submitted with delay or did not complete ? If there is at least one, what is the justification?
- Overall, considering the answers to the previous items, if your dedication were to be translated into attendance frequency, which frequency would it be ? [0% a 100%]

8. Evaluation aims to verify whether the objectives have been achieved or not. In this question, taking into consideration your initial expectations, make a self-evaluation about your performance. Address your ease with the contents, your understanding of the concepts, your assiduity to classes, your participation in classes, and your performance on the tasks. Feel free to add other comments.