

Final exam, ML (MDS + BDMA + MIRI)

21st of June, 2023

- In all your answers, conciseness and clarity will be most appreciated.
- All questions have equal weight.
- Exam duration: **2h**

Question 1

Describe the machine learning technique known as *bagging*. What are its key components? What are its main advantages?

Question 2

Describe *Bayes rule* in the context of classification. Give examples of classifiers based on this rule.

Question 3

Consider the following univariate regression dataset with three input examples (x_i, y_i) :

$$D = \{(1, 1), (2, 2), (3, 1)\}.$$

- a) Compute the mean square *loocv* error of ordinary linear regression.
- b) Compute the mean square training error of ordinary linear regression.

Question 4

Explain the effect of the following operations on bias/variance:

- a) Regularizing weights in a logistic regression model
- b) k in k -NN model
- c) pruning a decision tree
- d) increasing the number of hidden units in an MLP neural network

Question 5

You are a data analyst and you have to work on a binary classification problem. Describe resampling methodologies suitable in the following scenarios, include in each case the steps you would take in order to produce a final model:

- a) Data is very limited
- b) There is plenty of data and it is cheap to get new labelled data