First project:

You will have to form a group of 3 to 4 students.

Choose a dataset from the UC Irvine Machine Learning Repository (https://archive.ics.uci.edu/ml/index.php) with at least 5000 instances and 20 attributes for classification or regression. Compare how the different approaches seen in class perform on this dataset to predict accurately the classes or the values of the unlabeled data. You should determine what are the best hyper-parameters for each approach you are using. You could use any Python libraries.

Pages: 5 - 8 pages

- include a presentation of the research questions

- the chosen methods to tackle them

- a presentation of the results and discussion and a conclusion/future work

- attach a description of the participation of each student to the project