

TP Data Mining

Download the data set file in the Challenge section. Analyze the data and make a report with the obtained results. Note that even if classes are given in the file, we suppose that we don't have such information. Hence only unsupervised methods should be employed.

The report should contain at least:

- a first chapter that present the characteristics of the data set with basics statistics and plots,
- a second chapter on attribute selection/extraction. If it is not necessary for your data set, justify why,
- a third chapter on clustering results and explanation about the best results/best parameters.

Each result should be explained. Note that bad results can also be interesting to present if well argued.

The pages number is not important, just make sure to present every thing asked. Don't forget to add your name on the report and to give your python code and the version of python you used.