Machine Learning 2023

Project guidelines

The final grade is the grade of the project that splits into two parts: the report (incl. codes and supplementary material) counts for 70% and the presentation (incl. slides) counts for 30%. Grades are rounded to the closest 0.1. The final grade is rounded to the closest 0.5.

The project is conducted in groups of 1-2 students.

Originality and style:

The project consists of an original analysis of a data set. This means that re-analyses of existing data sets are possible when they are not just a repetition of the previous analysis. The re-analysis must have new elements (new models, features, outcome labels, approaches, interpretations, etc.). In any case, the data source must be given. If a previous analysis was done in the reference, this analysis must be explained, and the original aspects of the new analysis must be made clear.

Completeness:

Your analysis must match the complexity and uniqueness of your dataset; your analysis must be comprehensive and not leave major aspects of your dataset unaddressed. The project must include at least:

- Cleaning of the data, creation of new features if needed.
- Exploratory Data Analysis.
- Supervised learning analysis: two or more models, two or more scores, a tuning of one or more hyperparameters per model, a data splitting (if a training/test set split is enough for the global analysis, at least one CV or bootstrap must be used), interpretation of the model.
- Unsupervised analysis: clustering and/or dimension reduction.

Deliverables:

- General: a project proposal
- Report: a full PDF or HTML report of your analysis with reproducible code and figures, not exceeding 30 pages (or equivalent for HTML), including title page and appendices.
- Presentation: the slides of your presentation.

Deadlines:

You must meet the following deadlines (by 23h59):

- Project proposal: 3rd April (week 7)
- Project update: 8th May (week 11)
- Project report: 12th June (week 16 – 1st week of the exam session)
- Presentation slide: 18th June (Sunday before the final presentation)

Not meeting the deadline penalizes the grade by 0.1 per started hour of delay. Maximum penalty of 0.5 for the project proposal and up to 0.5 for the project update. No maximum penalty for the project report and presentation slide.

<u>Presentations:</u>

Presentations will be organized on site on Monday, June 19th in a 15+10-minute format (tentative).

If the number of groups is large, then either the presentation time will be reduced, or, if not possible, another oral presentation session will be organized during the June exam session.