

Data Science Lab

A short guide to laboratory submissions

Politecnico di Torino
November 6, 2019

Intro

This guide is intended for those of you that want to submit their laboratory work and participate to the weekly competitions. Even though laboratories submissions will not count on your final exam mark, we strongly encourage you to test your results. Doing so, you can practice with the same system that you will use for the final exam. At the same time, you will compete against your colleagues.

1 Rules and other information

Deadline. Each submission will have a strict deadline. Uploads later than the deadline will not be evaluated.

Evaluation and leaderboard. Your results will be evaluated on a specific metric. Since the goal changes from time to time, you can find which metric will be actually used in the laboratory text. Based on your score, you will be placed in the leaderboard.

We want you to provide a label for every point of the actual dataset. Nonetheless, at evaluation time, data are split in two chunks randomly: a *public* and a *private* one. Every time you perform a submission, the first chunk (i.e. the *public* one) is used to compute your temporary score and place you into the online leaderboard. In case of multiple submissions, only the best performing one will be retained and used to place you into the public leaderboard.

Once the competition is closed (i.e. right after the deadline) a new evaluation pass is run on the second, *private* chunk. This second evaluation will create a new, final leaderboard to declare the winners. Consider that the two chunks are sampled uniformly from the original dataset, hence they share the same properties and data distributions. Figure 1 shows a graphical representation for your submission pipeline.

Baseline. For each competition you will have a baseline score. This is a score achieved with a standard approach. You should try to achieve a better result than this value.

Personal key. To submit your result our system will ask for your personal key. Each student will receive its own personal key via email. These are strictly personal and must be used only in the submission form.

Submissions limit. In order to avoid overloading the server, you will have an upload limit (e.g. #upload per minute). Typically, we will give you this limit in the laboratory text and will be clearly highlighted in the platform.

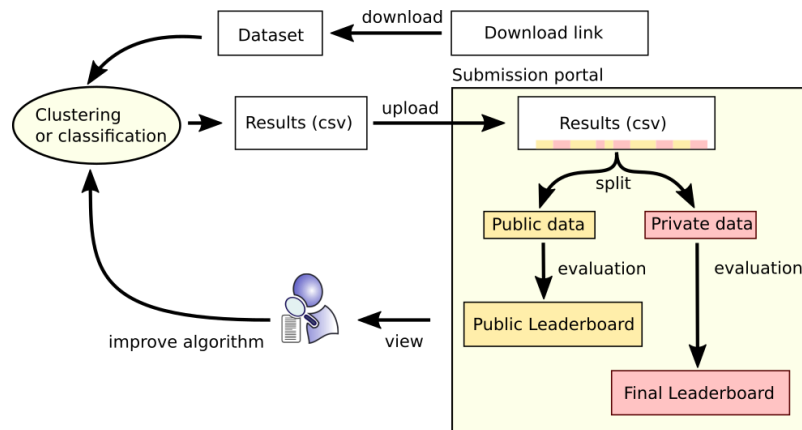


Figure 1: A graphical representation for the submission process