

Some news about the data challenge

Data challenge

- We will post the homework's results tomorrow;
- **Deadline for the data challenge is March 15th;**
- You need to provide a 2 pages pdf report on March 22nd, as well as the source code to reproduce your results; You need to have a simple script "start" that reproduces your submission file.
- As usual, send it to `julien.mairal@inria.fr`;
- **Any machine learning/signal processing/computer vision external library is forbidden. You need to become a maker, not a user.** Only optimization libraries (cvx, cplex, LP/QP solvers) are allowed.
- Do not wait for the last day to upload your results. You only get two shots every day.
- Half of the grade will be based on your score; 25% based on your report and 25% based on the methodology you used.

Data challenge and bugs

- All teams should target a **minimum 93% accuracy**; this is what is achieved by an SVM with Gaussian kernel and badly tuned parameters.
- If you do not achieve 93% accuracy with a reasonable kernel method, this means you have some **BUG**.

A bug can have many forms: you may

- incorrectly read the input data (are you truncating large and small values by mistake?);
- solve the wrong optimization problem; misread the labels (does the first label start to 0 or 1?).
- use only one part of the training data by mistake...