Challenge data 2020 for companies

challengedata.ens.fr

Data science challenges are organized by the Data team at ENS in order to put companies in contact with excellent students, engineers and researchers in data science. The challenges proposed by companies or laboratories arise from practical problems they encounter in their activity. They are organized in the form of machine learning competitions, through the platform challengedata.ens.fr (in a Kaggle fashion). More than 1800 students, researchers and engineers have participated to these challenges in 2019.

This initiative is in the spirit of scientific exchange, such that data and results be shared. The datasets made available by the companies and laboratories should be non-confidential. Algorithmic reports of students and researchers can be accessed by the companies. The web platform supports the data exchange and the automatic scoring of the results of the participants. The evolution of the scores and rankings can be monitored in real time.

Why participate?

These challenges offer companies an access to state-of-the-art algorithms, whatever the proposed problem may be, thanks to the live ranking of the results and the diversity of the participants. Companies can foster links with the best students and professionals participating in their challenge. The link can be established through the presentation of the challenges in the prestigious setting of Collège de France, through online interaction during the challenge, and through the closing ceremony, also at Collège de France. Over the last years, these links have given rise to numerous internships and hirings.
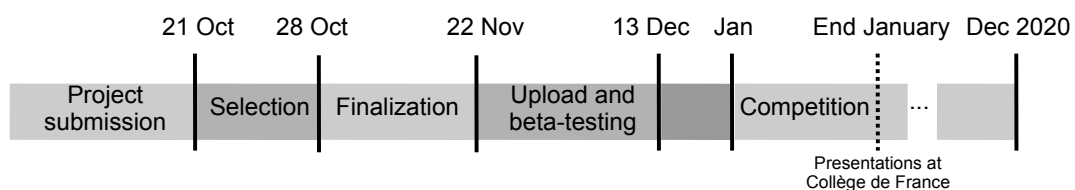
What types of projects?

Each company can propose a supervised learning problem (classification, regression, prediction, ranking...) by providing training and test datasets. All data types are accepted, be them medical signals, images, videos, physical measures, sounds, financial time series, texts, marketing questionnaires, web click data, etc. It is recommended to define a problem which can be easily understood by participants and the solution of which has a true impact, even if it can be difficult and involve sophisticated algorithms.

How to participate?

Interested companies are invited to submit a project as soon as possible by October 21st 2019, by email at challenge DOT data AT ens DOT fr. The document should be no more than two pages long, in English only, and include the following items:

- the name of the company and a short description of its activities;
- the name and email address of the person responsible for the project;
- the definition of the machine learning project and its interest, specifying if it is a classification, regression, prediction, ranking, etc, task;
- a description of the input variables (x) and the output variables (y), specifying the number of examples, split between training and test;
- a metric allowing to quantify prediction errors on y;
- a proposition of a single benchmark algorithm to obtain first prediction results on the dataset (which will be implemented).

Proposals shall follow the template joined.

21 Oct	28 Oct		22 Nov	13 Dec	Jan	End January	Dec 2020
Project submission	Selection	Finalization	Upload and beta-testing		Competition	⋯	

Presentations at Collège de France

An initiative supported by