

RAMP: collaborative data challenges run by the Paris-Saclay Center for Data Science

jeudi 6 octobre 2016 11:45 (35 minutes)

RAMP (Rapid Analytics and Model Prototyping) are collaborative data challenges, which aim at connecting data science to domain sciences, which is the main mission of the Paris-Saclay Center for Data Science.

The lifecycle of a RAMP is the following: A domain science expert brings data and an associated prediction problem, and a data scientist helps clean the data and formulate a machine learning problem. A RAMP event is then organized with 30-50 data scientists and domain scientists. During the RAMP, the participants submit a model generating predictive solutions. The models are trained on our back-end. The scores are displayed on a leaderboard. All participants have access to all code, and they are encouraged to look at and to reuse each other's solutions.

Past RAMP lead to a significant improvement over the baseline. Since the organizers have access to all the code, the result of the event is a fully functioning near-optimal prototype.

In this presentation, we are introducing the two tools we are building to run these collaborative data challenges:

- databoard: a platform where participants can submit their code (machine learning models), see code from other participants, and follow the leaderboard.
- datarun: a platform where submitted models are trained.

Orateur: Mlle MARINI, Camille

Classification de Session: PlaScido