

What is wrong with data challenges

Friday, 11 December 2015 14:00 (40 minutes)

We will develop a constructive criticism of the data challenge format practiced today. It will be illustrated by our story of the HiggsML challenge, but our conclusions will go beyond. In a nutshell, challenges are long job interviews for participants, publicity for organizers, and benchmarking and teaching aids for the data science community. What are they not? They will not deliver a workable solution to your problem, not even a prototype, partly because the very problem you can squeeze into the competitive gaming mechanism is a diluted or abstract version of the real problem you want to solve. You will have no access to the data scientists participating in the challenge, unless of course you can hire them. They incentivize neither collaboration nor creativity.

In the last third of the talk I will describe the format and tool that we have been developing to run collaborative hackatons (RAMPs for Rapid Analytics and Model Prototyping) at the Paris-Saclay Center for Data Science which implements some of the missing features of the classical challenge format.

Presenter: KÉGL, Balázs (CNRS / Université Paris-Saclay)