

---

# Design and Analysis of Experiments: A Challenge Approach in Teaching

---

A. Pavao<sup>1</sup>(\*)    D. Kalainathan<sup>1</sup>    L. Sun-Hosoya<sup>1</sup>    K. Bennett<sup>2</sup>    I. Guyon<sup>1</sup>

(\*) **Corresponding author:** adrien.pavao@gmail.com

1. Université Paris-Saclay (UPSud/INRIA), Gif-sur-Yvette, France

2. Rensselaer Polytechnic Institute, Troy, NY, USA

## Abstract

Over the past few years, we have explored the benefits of involving students both in organizing and in participating in challenges as a pedagogical tool, as part of an international collaboration. Engaging in the design and resolution of a competition can be seen as a hands-on means of learning proper design and analysis of experiments and gaining a deeper understanding other aspects of Machine Learning. **Graduate students** of University Paris-Sud (Paris, France) are involved in class projects in **creating a challenge end-to-end**, from defining the research problem, collecting or formatting data, creating a starting kit, to implementing and testing the website. The application domains and types of data are extremely diverse: medicine, ecology, marketing, computer vision, recommendation, text processing, etc. The challenges thus created are then **used as class projects of undergraduate students** who have to solve them, both at University Paris-Sud, and at Rensselaer Polytechnic Institute (RPI, New York, USA), to provide rich learning experiences at scale. New this year, students are involved in creating challenges motivated by “AI for good” and will create re-usable templates to inspire others to create challenges for the benefit of humanity.

## 1 Pedagogical motivation

Over the past few years, challenges have become pervasive in teaching. For instance, Kaggle “in class” has hundreds of competitions<sup>1</sup>. The benefit of using competitions in teaching includes motivating students and facilitating grading. However, little use has been made so far of involving students in the **design and implementation of competitions**. Clearly this is a more difficult task and the most sophisticated challenges can take several years of maturation and the involvement of mature researchers. However, relatively simple competitions, of a level of difficulty that can be used to train undergraduate students, can easily be designed and implemented by graduate students, as part of class projects (typically classification and regression problems, also occasionally recommendation and reinforcement learning problems have been addressed). This allows them to gain hand-on practice of experimental design and harness the difficulties raised by defining well tasks and metrics, collecting and preparing data ensuring that there are enough samples and no bias or data leakage is present, preparing baseline methods, and presenting their challenge to engage as many participants as possible to solve it. In this process, students also learn to work in teams, meet strict deadlines, and acquire programming skills in Python, including to master toolkits such as scikit-learn and Keras, and good coding practices including version control with Github and the use of dockers. Emphasis is put on creating a fully working end-to-end “product” (a challenge), which will then be used by real “customers” (the undergraduate students). Quality of communication is also stressed by making the

---

<sup>1</sup><https://www.kaggle.com/competitions?sortBy=grouped&group=inClass&page=1&pageSize=20>

graduate students produce a short advertising video and presenting their challenge in class to the undergraduate student, who get to choose one of them for their project. We have been conducting this type of educational program since 2016<sup>2</sup>. Each year 30 to 40 graduate student create challenges as part of their master program in data science and about 100 second year undergraduate students solve them over a 12-week project period. Thus we have already trained over 500 students with this program. We have recruited several alumni as challenge co-organizers for larger research challenges, which have been selected as part of the NeurIPS competition program, such as the TrackML particle physics challenge<sup>3</sup> and the AutoDL challenge<sup>4</sup>.

## 2 Community impact

Engaging graduate students in the design of challenges has important far reaching impact. With the current rapid growth of AI research and applications, there are both unprecedented opportunities and legitimate worries about its potential misuses. In this context, it is important to raise awareness among students of good data science methodology with respect to study design and modeling. Recognizing that there is no good data science without good data, we want to educate them to conduct proper data collection and preparation. Our objective is to instill good practices to reduce problems resulting from bias in data or irreproducible results due to lack of data. We also encourage the protection of data confidentiality or privacy by making use of software replacing real data by realistic synthetic data. This facilitates broadening access to undergraduate students to confidential or private data having a commercial value or the potential to harm individuals.

New this year, one original complementary aspect of our educational projects is to turn the challenges that have been designed into simplified templates, using place holder data (e.g. synthetic data, as described above), and including ready-made starter solutions. Such templates will showcase a wide variety of data-driven AI applications, to trigger the imagination of researchers world-wide, with no particular AI expertise. By simply cloning a template and replacing the data, an organisation could get immediate baseline results and eventually refine them by opening the challenge as an internal or external competition. To facilitate this process, we are making available for free our open-source challenge platform Codalab<sup>5</sup> and will provide extra computational resources on the platform, based on merit and need of the challenge organizers.

Besides facilitating the re-use of our challenges by other instructors, disseminating our challenge templates will also benefit low budget entrepreneurs eager to bring AI solutions to new domains. Finally, volunteers who want to contribute to “AI for good” will have a platform to quickly put together applications by cloning challenge templates. We believe that this will be an important contribution to the democratization of AI.

## 3 Conclusion

Teaching students to organize and participate in machine learning competitions prepares them well to become data scientists and at the same time helps us grow a community of challenge organizers with good practices and disseminate challenge templates, which can serve to spread of AI for all and AI for good.

### Acknowledgements

Many people participated to this project. We are very grateful to Eric Carmichael and Tyler Thomas for developing and maintaining the challenge platform Codalab and the teaching tools ChaLab and Chagrade and to the teaching assistants Zhengying Liu and Balthazar Donon as well as class mentors including Magali Richard, Guillaume Charpiat, and Antoine Marot. The project was funded in part by ChaLearn, Université Paris-Saclay “big data” chair of Isabelle Guyon, the EU project HADACA (EIT Health) and the United Health Foundation (INCITE project, RPI, New York).

---

<sup>2</sup><http://saclay.chalearn.org/>

<sup>3</sup><https://sites.google.com/site/trackmlparticle/>

<sup>4</sup><https://autodl.chalearn.org/>

<sup>5</sup><http://competitions.codalab.org>