Competitions and challenges for education and continuous learning

Kristin P. Bennett BENNEK@RPI.EDU
Rensselaer Polytechnic Institute, USA

Isabelle Guyon GUYON@CHALEARN.ORG
Université Paris-Saclay, France, ChaLearn, USA, and Google, USA

Adrien Pavão PAVAO@MLCHALLENGES.COM
Université Paris-Saclay, France

Evelyne Viegas EVELYNEV@MICROSOFT.COM
Microsoft, USA

Reviewed on OpenReview: `https://openreview.net/forum?id=XXXX`

Abstract

Involving students and trainees both in organizing and in participating in competitions and challenges is a powerful pedagogical tool. From kindergarten through the last years of graduate studies, or as a way to upskill to gain new skills on the job, competitions can gamify the learning process and thus motivate people to explore by themselves, assimilate a variety of material, and expand their capabilities. Competitions also contribute to engaging students in learning about problems of societal importance. At the graduate level, designing and implementing competitions can be seen as a hands-on means of learning proper design and analysis of experiments. In the workplace, this remains an effective way to embrace continuous learning and remain up to date as technologies, tools, and approaches evolve. In this chapter, we present a synthesis of various hands-on teaching and learning experiences using competitions as a medium. We report educational efforts conducted in kindergarten, high school, university and working environment. For younger students, competitions take more the form of individual projects qualitatively evaluated by a jury, while at the university level they take the form of automatically graded homework addressing a research problem. The application domains can be very diverse: medicine, ecology, marketing, computer vision, recommendation, text processing, etc., and students enjoy being involved in creating challenges motivated by humanitarian purposes. Within a professional or working environment, competitions are used to help with upskilling, learning new approaches, and tools that can directly be used in one's job or to update one's resume to remain relevant and competitive in the market place.

Keywords: education, teaching, continuous learning

1 Introduction

Competitions and challenges (together referred to as skill-based contests) are a form of "serious games". Already in 2017, the NeurIPS workshop on "Challenges in Machine Learning" (CiML) focused on gaming and education made the connection between challenges and games. Since then, hundreds of challenges in Machine Learning and AI have been organized on platforms such a Kaggle community competitions Goldbloom and Hamner (2010) and RAMP Kégl et al. (2018), as well as Codalab competitions Pavao et al. (2022).

According to Nguyen (2021): "Research shows that using games in teaching can help increase student participation, foster social and emotional learning, and motivate students to take risks."

Engaging in games or competitions as part of an educational program provides a more collaborative and engaging experience, particularly if team work is encouraged, for students who struggle with passive learning, heavy on lectures and book reading. During Covid times, teaching via participating (or organizing) challenges helped students being connected and motivated remotely, by working on a project together. These claims are supported by research about using games in teaching Plass et al. (2015); Wu et al. (2012); Ifenthaler et al. (2012), which presumably helps increase student participation, foster social and emotional learning Hromek and Roffey (2009), and motivate students to take risks.

While the terminology "competition" and "challenge" is often used interchangeably, in this chapter, we use the following definitions. We consider only scientific evaluations running over a finite amount of time, referred to as "skill-based contests" (as opposed to benchmarks, which are run on an on-going basis and "chance contests" in which chance plays a role in winning). We call **"challenge"** those skilled-based contests in which participants are called to solve **organizer-specified tasks** evaluated by given metrics. Challenge entries are usually results of predictions made on a challenge test set or code to perform a specific task. Short challenges are sometimes referred to as **"hackathons"**. In contrast, **"competitions"** are more open-ended. The participants can be called to **define and solve their own problem**. Competition entries can include reports, prototypes, demonstrations, live performances, presentations, etc. Competitions are often judged by a jury on the basis of both quantitative and qualitative metrics, while challenge evaluations are usually automated and obtained by computing scores according to pre-defined metrics that are posted on a leaderboard.

Neither competitions nor challenges are not substitutes for other forms of learning. Like any educational tool, they need to be well-planned and integrated only when they are relevant to the learning objectives. In this chapter, we explore various aspects of including competitions or challenges as part of curricula, as participants but also as organizers, then analyze various case studies.

2 Students entering a competition

A seemingly easy way to engage students at low preparation investment is to enroll them in a competition organized by a third party. This can constitute a **project-based class** giving a lot of freedom of creativity to students.

At the K-12 level, that is between kindergarden and last year before college or university, Technovation offers programs to engage underrepresented groups, particularly young women (ages 8-18) to become leaders, creators and problem-solvers. The competitors form teams coached by parents and educators to solve real world problems, such as finding safe drinking water, identifying and removing invasive species; monitoring air quality, etc. Technovation organizes an annual competition, through which participants identify problems in their communities and use mobile and AI technologies to develop solutions. Technovation has been in operation for 14 years, reaching more than 160,000 participants through its competitions. See Section 6 for more details.

At the high-school level, robotics competitions offer an opportunity the break the ice with intimidating technology. Well known competitions include the "First Lego League" and the "Sumo Robot League". Then under-graduate level competitions include Duckietown, which offers each year several AI Driving Olympics competitions (AI-DO) for small self-driving robots, which can be purchased at a low price and self-assembled. The Duckietown project[1] was conceived in 2016

1. https://duckietown.com/ai-do-at-neurips-is-over-congratulations-to-our-winners/

as a graduate class at MIT. The Duckietown Foundation debuted AI-DO in December 2018 at the NeurIPS conference. The platform is used at several universities around the globe, including NCTU in Taiwan, Tsinghua in China, and RPI in the United States. Another well-known and well established competition is RoboCup, which offers junior level entry leagues, all the way to advanced professional leagues.

At the graduate level, students can directly enter "high profile" international competitions. This can be very motivating, as the prize, which can reach thousands to hundreds of thousands of dollars, can help funding their PhD. research and beyond. For example, the XPrize has been proposing numerous bold challenges, from sending a rocket to the moon to providing solutions to the Covid crisis. Recently, they have been attracting a lot of attention with the XPrize Carbon Removal, aimed at fighting climate change and rebalancing Earth's carbon cycle. Funded by Elon Musk and the Musk Foundation, this $100M competition is its largest incentive prize. For the first milestone, several student teams were awarded 100,000 dollars to pursue their research, including women-led U. Miami student team for their proposal of ocean-based carbon removal, and women-led U. Washington student team for their novel carbon soil gas monitoring sensors.

3 Students entering a challenge

Compare to competitions, challenges are a more contrained format of skill-based contests. They usually organized on a challenge platform, such a Kaggle community competitions, RAMP, or Codalab competitions. Many machine learning, natural language processing, and computer vision conferences organize challenges every year, so it is not difficult to find a suitable challenge to teach a class. For example, in 2022, the NeurIPS conference offered 25 high-end peer reviewed challenges.

The benefit of using challenges in teaching includes motivating students and facilitating grading (since the leaderboard scores can readily be used as part of the grade). In our experience, it is important to have a very structured class to set the student expectations, and use the competition or challenge as a medium (a means to an end), avoiding to emphasize winning as the essential goal. This last point is facilitated by grading the students on several aspects other than winning (e.g. quizzes, oral presentation, written report) and organizing students in **teams** and/or **leagues**. Teams are groups of students working together towards solving a problem making challenge entries; leagues are sets of teams, generally of a similar level or interest, competing with one another. See Section 5 for more recommendations on grading.

For project-based classes (or for the final project of a class), one may consider **letting student choose their own challenge** and just deliver a report (possibly assorted with in-class presentations or a poster session). The instructor may want to narrow down choices for multiple reasons: finding a good challenge is time consuming, and students are not necessarily the best judges of what a good challenge is. They can choose a challenge, which is either too easy or too hard, or which does not offer good learning opportunities. The choice of challenges does not necessarily need to include only on-going challenges. In fact, entering a challenge, which is already over, presents multiple benefits: lowering the barrier of entry (with available solutions already published), and diminishing the importance of winning (in favor of learning). However, the students must then highlight clearly, in their report, their personal contribution.

Another way of incorporating challenges in a class is to design homework as challenges, each assignment being delivered as an entry in a mini-challenge. In that case, the instructor(s) design the challenge. Examples of such challenges include:

- Iris data challenge, organized to a beginner undergraduate class at U. Paris-Saclay, 105 participants.
- Artificial Neural Networks and Deep Learning 2021, organized by Politecnico Milano, 486 participants.
- Linear model contest, organized by U. Washington, 322 participants.
- Black Box Optimization challenge, organized by AI master optimization class at U. Paris-Saclay, 29 participants. A simplified remake of NeurIPS 2020 BBO challenge.
- Evolutional Reinforcement Learning, organized by AI master RL class at U. Paris-Saclay, 29 participants.
- Prediction of mortality given medical records, organized repeatedly by RPI used in data science class taught for serveral years with 50 to 91 participants each semester. See case study.
- ChemsRUs, organized repeatedly by RPI for a data science class (predict biodegradability of molecules) with 50 to 91 participants.See case study 8.

Some of these challenges are re-makes of large international challenges, which have been simplified. Obviously, having good introductory material, e.g. in the form of a R-notebook or a Jupyter-notebook, is essential. Also important: one must avoid wasting student's time with complex installation instructions. Thus, it is advisable to rely on ironed-out tools, such as:

- Anaconda Python installation;
- Google Colab on-line Jupyter notebooks;
- Scikit-learn ML library Pedregosa et al. (2011);
- PyTorch Deep Learning framework Paszke et al. (2017).
- R Project for Statistical Learning R Core Team (2021).

Challenge-based classes can also be an opportunity to give students good habits, such as using revision control (Git and Github), learning about Dockers, etc.

It can also be very motivating for students to participate in challenges designed by other students. See Section 4.

4 Students organizing a challenge

Université Paris-Saclay offers each year a class on creation of AI challenges, which are then solved by other students (see a list of past challenges organized by students) as part of their class requirements. This approach is illustrated by Figure 1.

While it has become mainstream to let students enter competitions or challenges, as part of class projects or homework, little has been do so far to involve students in the **design and implementation of competitions** or challenges. Clearly this is a more difficult task. Indeed, sophisticated challenges can take several months or years of maturation, and the involvement of many researchers.

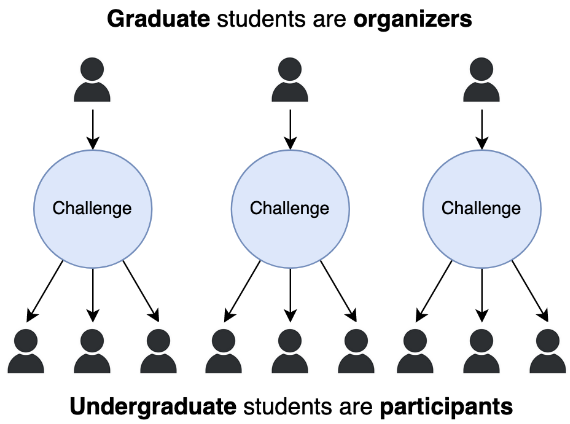

Figure 1: Graduate students organize challenges in which undergraduate students participate.

However, relatively simple challenges, 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**, but eventually recommendation or reinforcement learning problems). This allows them to gain **hands-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,
- Ensuring that no bias or data leakage is present in data,
- Preparing baseline methods.

In this process, students also learn about:

- Working in teams,
- Meeting strict deadlines,
- Acquiring good programming skills (including programming e.g. in Python or R and mastering toolkits such as scikit-learn and Keras, and learning about Github and Dockers),
- Preparing good didactic material (starting kit),
- Presenting their work (orally and in a written report),
- Promoting their challenge to engage as many participants as possible.

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 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.

This type of educational program has been taking place since 2016 at Université Paris-Saclay Pavao et al. (2019) (see the 2021/2022 edition). Each year 30 to 40 graduate students 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. This program has also used student-designed challenges as master-level projects. Already 1000 students have been trained through this program. Several alumni have become co-organizers of larger research challenges, which have been selected as part of the NeurIPS competition program, such as the TrackML particle physics challenge Calafiura et al. (2018), the AutoDL challenge series Liu et al. (2020), the Meta-Learning challenge series Carrión-Ojeda et al. (2022), and the reinforcement learning Learning to Run a Power Network (L2RPN) challenge series Marot et al. (2021, 2019).

Engaging graduate students in the design of challenges has an 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 familiarize students with good data science methodology, with respect to study design and modeling. Recognizing that there is no good data science without good data, it is important to educate students to conduct proper data collection and preparation. The objective should be to instill them good practices to reduce problems resulting from bias in data or non reproducible results due to lack of data. At RPI, the student challenge program has also been encouraging the protection of **data confidentiality or privacy** by replacing real data by realistic synthetic data Yale et al. (2020). This facilitates broadening access to undergraduate students to confidential or private data having a commercial value or the potential to harm individuals. Awareness should also be raised to issues related to **ethics and fairness**. To that end, Université Paris-Saclay offers a class on Fairness in Artificial Intelligence, in parallel with the challenge class.

From previously designed challenges constitute progressively a "library of challenge designs"[2]). Previously designed challenges can be cloned and serve as templates for new challenges. See chapter 10 for a tutorial on how to easily clone CodaLab challenges and create your own challenge or design a simplified template for your students as a starting point.

5 Evaluating teaching effectiveness and measuring impact

In this section, we provide a few tips on how to grade students and evaluate their learning experience. We note that using the raw scores on challenge leaderboards is not necessarily a good means of evaluating students. However, challenges can provide valuable and effective pedagogical tools for evaluation for busy teachers. Challenges offer a unique opportunity to measure such effectiveness, by monitoring engagement (number of submissions), collaborations (code sharing), good programming habits (e.g. good use of GitHub). Additionally, students can be asked to complete quizzes or surveys before, during, and after the course program.

Entry survey

Before the course begins, we (Isabelle Guyon and Adrien Pavão) usually ask students to complete an "entry survey", to check their level, interests, and level of motivation. The answers serve to form teams and assign students to a challenge corresponding both to their wishes and their skill level. We advocate that this survey should not be graded, but rewarded by a flat number of "bonus points" for

2. `https://saclay.chalearn.org/`

completing them. The students can complete the survey at home and are encouraged to search on the Internet to answer the questions, but answer them in their own words.

We use a similar survey regardless whether the class is about organizing a challenge or participating in a challenge. Typical questions we ask include:

- What is your level in Python programming?
- Have you already participated in a machine learning or data science challenge? If so, describe your experience and provide the URL.
- Explain the difference between supervised and unsupervised learning.
- Explain the difference between classification and regression.
- Explain what is "cross-validation".
- Explain the difference between aleatoric and epistemic uncertainty.
- Explain what "data leakage" is. Give an example.
- What machine learning toolkit do you prefer and why (Scikit-learn, Tensorflow, Keras, Pytorch, other)? Justify your choice.
- What to do think is hardest to deal with: too little or too much data? Justify your answer.
- What application domains are of interest to you?
  - Biology and medicine.
  - Ecology, energy and sustainability.
  - Internet, social media, and advertising.
  - Market analysis and financial data.
  - Image, audio, speech, video, and other sensor data.
  - Text processing, language understanding.
  - Physics and chemistry.
  - Ethics, fairness, and privacy.
  - Engineering, manufacturing, quality assurance.
  - Robotics and control.
  - Education.
  - Sports data analysis/prediction.
  - Games (Chess, Go, ...).
  - Generative adversarial networks.
- You will work as a team. Characterize your skills:
  - Good programmer.
  - Good sense of user interfaces.

- Good artistic sense.
- Well organized.
- Capable of coordinating a team.
- Good written English.
- Good oral English.
- Good practical experience of machine learning
- Good in statistics or learning theory.

It is impressive how efficient such questions are to evaluate, not only the level of the students, but their motivation and willingness to learn, and their capability to look for answers by themselves and assimilate them. Team leaders are chosen on the basis of self-declared capacity of leading a team, being well organized, and diligently trying their best to provide good answers. The rest of the students are generally grouped on the basis of topic affinity. Mixing weak and strong students in the same team is usually ineffective, because the stronger students end up doing everything and ignoring the weak students. Hence we tend to also regroup students by strength, as a secondary criterion after topic affinity.

In our challenge classes at Université Paris-Saclay, we group students in teams of five or six people. The teams are made by the professor, based on the survey answers. The team members can then elect to dispatch among themselves complementary roles and eventually work in pairs: (1) Coordinator (will be responsible to submit all the homework in time); (2) Architect (will oversee the end-to-end "product" design); (3) Domain expert (will oversee the data preparation and task definition); (4) ML expert (will oversee the preparation of the baseline software); (5) Data analyst (will oversee the production of interesting results); (6) Test engineer (will be responsible to make sure that everything works). The assignments of roles to people are flexible (more than one role per person and more than one person per role are possible).

Milestones and deliverables

For a challenge class to be successful, it is important to give the student, in advance, a clear schedule of milestones and deliverable. Depending on whether the class is a challenge design class or a challenge participation class, those are slightly different, but here are some common final deliverables:

- **Proposal:** the students must describe in a few pages what they plan to do.
- **Video:** a mini 3 minutes "promotional video", inspired by competitions of "my thesis in 180 seconds". This teaches them how to communicate succinctly and how to make a video.
- **Presentation:** a short 10-15 minute in-class presentation. This makes them practice public speaking and explaining technical contents.
- **Report:** a short 6-8 page techreport. This teaches them how to write a scientific paper with systematic comparisons of methods, graphs, and error bars.

For the challenge preparation class, other deliverable include:

- **Starting kit:** a Jupyter notebook serving as tutorial material to prepare and entry into the challenge, including sample data and a baseline method.

- **Website:** a working implementation of their challenge on a platform.

- **Tests:** a series of tests checking their final product.

For the challenge participation class, we usually intermix in-class practical work (based on Jupyter notebooks) to be finished at home as homework, quizzes, and code reviews. The ranking into the challenge weighs little in the final grade.

For details, see for example the Syllabus of the AI challenge creation class (2021/2022), and the Syllabus of the Data Science challenge class 2019/2020 at Université Paris-Saclay.

To keep the students engaged and motivated, all homework and deliverables have strict deadlines. However, if delivered on time, the instructors give only a temporary grade, which can be improved by submitting a second corrected version. This "second chance" method motivates students to work diligently and efficiently in the classes. Since they clearly know how to improve their grades from the instructor's reviews, the students usually work hard on the second versions.

Class evaluation

It is important to evaluate how much students assimilate the material and if they are encouraged to continue pursuing a career in data science or artificial intelligence. A post-program survey can help collecting such information.

For example, Rensselaer Polytechnic Institute (RPI0 formally evaluated it's low-barrier pipeline into applied undergraduate research consisting of an introductory data analytics course called Introduction to Data Mathematics Course (Case study 2 below), followed by a course-based undergraduate research experience (Case study 3 below) Bennett et al. (2022). The responses from 118 students, presented in Figure 2, show that the majority of respondents (80%) might or do plan to pursue additional courses or experiences relevant to data analytics. Almost half of the respondents plan to pursue Internships/Coops Related to Data Analytics (46%). Almost all respondents (93%) agreed or strongly agreed that as a result of taking the course they want to improve their computer skills. The great majority also recognized the value of data analytics to the healthcare field (87%), understand that data analytics will be of value regardless of their career path (94%), and aspire to use data to solve real world problems (85%). Most respondents want to learn ways to apply data analytics in other areas (90%) and the majority (67%), want to have a career in data analytics, almost double the previous response. Over three-quarters of respondents agreed or strongly agreed that they want to continue doing research projects that involve data analytics (80%).

Women are a third of students in these courses, which is proportionate to their representation at the Institute. The four most common majors of students are Mathematics (37%), Computer Science (22%), Biochemistry and Biophysics (15%), and biology (70%). A third of the undergraduates taking these courses are dual majors. The most common second major for almost half of the dual major students was mathematics (49%) followed by computer science (19%).

Measuring impact

Much remains to be done in terms of measuring impact. Nalia Kabeer presents a seminal framework for measuring empowerment, measuring increases in:

- **Resources:** Increased access to material, human, and social resources.

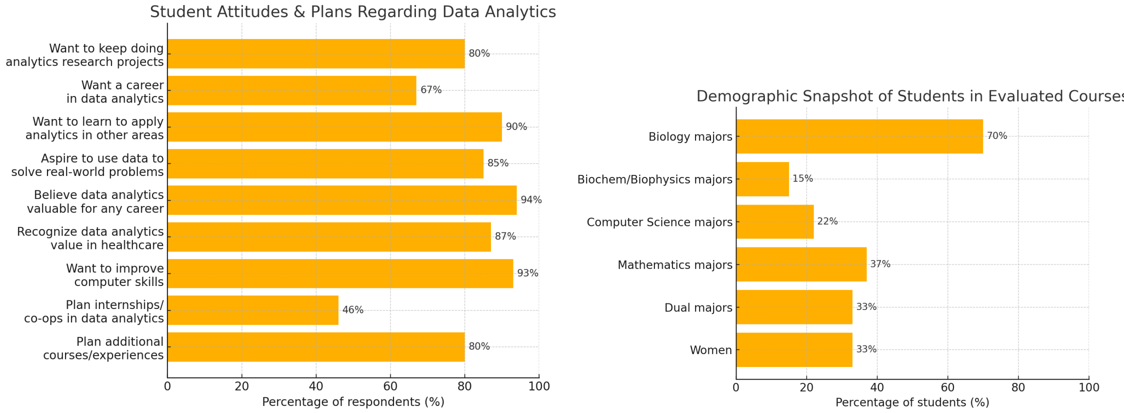

Figure 2: **Left:** Student attitudes and future plans related to data analytics after completing the evaluated courses. **Right:** Demographic snapshot of students enrolled in the evaluated courses.

- **Agency:** Increased abilities, participation, voice, and influence in the family, workplace, school, community.
- **Achievements:** Meaningful improvements in well-being and life outcomes that result from increasing agency and education (Kabeer, 1999).

We devote the rest of this chapter to describing three case studies of uses of challenges or competitions in education.

6 Case study 1: K-12 competitions

This section covers engaging school-age children at grade levels K-12 in competitions organized by Technovation. The purpose of running competitions in K-12 is to broaden participation, and building skills, particularly debunking myths around competitions and gender. How to run successful competitions at the K-12 level that broaden participation is rooted in Bandura's Motivation Theory.

Technovation puts forward two research frameworks:

- Higher dosage and practice so learners move from "situational interest" - participating in a program because it is offered in their community - to "well-developed, individual interest" - where the activity becomes a core part of their identity Hidi and Renninger (2006).
- Adaptation of Bandura's self-efficacy theory Bandura (1997), which outlines four pillars present in every successful human behavioral intervention:
    - Exposure to mentors and stories modeling lifelong learning.
    - Multi-exposure learning experiences that are authentic, engaging and meaningful.
    - Supportive cheerleaders who hold learners to high expectations while providing necessary support.

- High-energy, dramatic, suspenseful social gatherings/competitions that help the community feel collective pride (and adrenaline) at their accomplishments.

Technovation is a technology education nonprofit with a mission to empower vulnerable groups (especially girls and women) to create technology-based solutions to problems in their communities. Over the past 14 years, Technovation has engaged 50,000 mentors and educators to support more than 250,000 participants across 100+ countries to tackle pressing problems ranging from climate change to substance abuse—and most recently, COVID-19 (Technovation impact report, 2020).

Technovation organizes yearly innovation competitions to solve community problems, with a combination of inexpensive hardware and software, typically web apps or smart phone apps. The curriculum includes coding tutorials that help students apply key programming principles to real-world problems. Such tutorials include lessons on how to access and use databases (which will help teams who develop apps that help communities better distribute resources) as well as lessons on how to integrate maps and location-based data into student projects. Most real-world problems, like COVID-19 or the climate crisis, are complex and ill-defined, operate at multiple scales across different disciplines in dynamic ways, and may not have a clear end. To prepare young people to face these challenges, Technovation created a "Solve-It" video series and an associated checklist for educators, based on Donna Meadows' primer on Thinking in Systems.

One program is called "Technovation Girls". The 2020 season of brought together **20,388 girls from 62 countries**, who, with the support of 10,491 mentors, educators, and chapter ambassadors, designed a total of 1,520 tech-based apps addressing problems in their communities. Problems ranged from environmental protection to gender-based violence, to COVID-19. For example, Technovation worked with an attorney from Hogan Lovells to apply the "Solve-It" checklist to a real example of a complex problem Technovation Girls teams frequently address—domestic violence. The video walked girls through the process of brainstorming and developing solutions that can help protect survivors of domestic abuse by exploring different potential effects of the solution on those who use it. This helped students consider the different systems that domestic violence is situated within, rather than approaching it as a stand-alone issue.

Every year, Technovation invites teams of girls around the world to learn and apply the skills needed to solve real-world problems through technology, during a 12-week program. According to Technovation, after participating in "Technovation Girls", students express a greater interest in technology and leadership, and 58% of the alumni enroll in more computer science courses. Alumni go on to start their own businesses, present at prestigious events, meet world leaders, and even return to support the next cohorts of Technovation Girls. Over the last 15 years, Technovation has trained 150,000 young women to be technology entrepreneurs and innovators, empowering them to solve problems in their communities using technology. Five or more years after participating in Technovation programs, alumnae still credit the program with influencing their interests, career, and professional pursuits. Impact indicators include that 76% of Technovation alumnae pursue a STEM degree compared to 21% of female undergraduates who received STEM-related degrees in 2012; 60% of Technovation alumnae work in STEM-related positions, compared to the 29% national average in 2013; 50% of alumnae are leading change in their communities. Mentors play a critical role: Over 90% of teams who successfully completed the program had a mentor, and 70% of the Technovation girls who completed the post survey shared that they were helped a lot by their mentors. These statistics are summarized in Figure 3

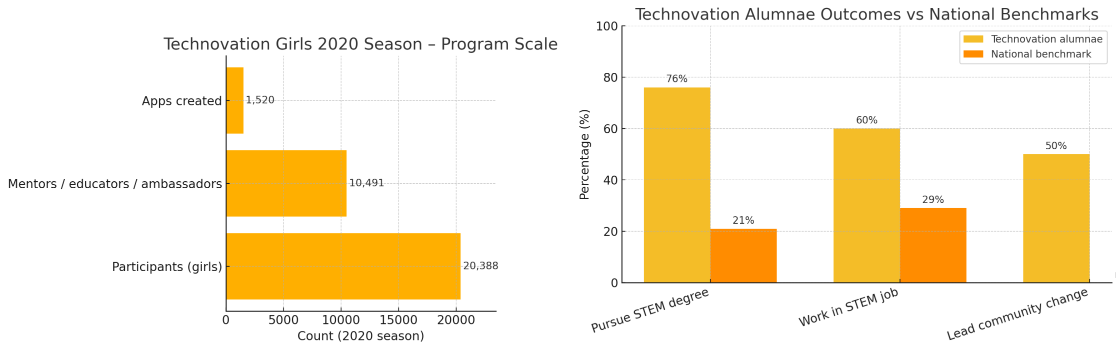

Figure 3: **Left:** Reach of the Technovation Girls 2020 season. **Right:** Long-term impact of Technovation participation.

Another program of Technovation engages families Chklovski et al. (2021). Over the course of 2 years it has engaged nearly 20,000 under-resourced $3^{rd} - 8^{th}$ grade students, parents and educators from 13 countries in a multi-week AI competition. Families worked together with the help of educators to identify meaningful problems in their communities and developed AI-prototypes to address them. The resulting projects (prototypes) were judged by a panel, using well-defined criteria. The program identified a high level of interest in underserved communities to develop and apply AI-literacy skills. Students followed 10 classes, and were tested their understanding of concepts through selected response questions on the curriculum platform. If they selected the wrong answer, they were prompted to try again. The competition element of the program provided the usual combination of pros and cons: time-based deadline that motivated families to persist and submit their prototypes, excitement of competing at a global level counterbalanced by stress, frustration, impatience, and forced deliberation. Strategies to improve retention include providing a variety of project-based learning lessons, starting from hands-on, unplugged activities and then moving onto software projects. Patience and commitment are needed: It takes 3 to 5 years to iteratively develop fun, engaging, effective curriculum, training and scalable program delivery methods. This level of patience and commitment is needed from all community and industry partners and funders.

7 Case study 2: Short-Competitions and Hackathons in the classroom

Short-term Competitions (over weeks) and Hackathons (over days) have been used in several contexts: as part of classes, student-organized events, conference tutorials, and summer schools. We provide a few hands-one experiences in this section.

Hackathons and competitions are used directly as part of **class activities**. RPI regularly uses two short-term challenges as assignments in its Intro to Data Mathematics (IDM) course. The challenges use project-based learning Anazifa and Djukri (2017) to encourage students to become creative data analysts by asking their own questions and developing their own strategies for addressing the challenge questions. This project-based learning strategy affords students the opportunity to develop DA knowledge and skills in context. We (Kristin Bennett, Adrien Pavao, and Isabelle Guyon) created two online challenges. In the first, To be or not to be: ICU Mortality, students predict whether an intensive care unit (ICU) patient lives or dies. ICU Mortality is a straight forward classification competition. Students do the challenge in teams as a two week assignment to practice their classification skills. They learn how to enter a challenge in preparation for the second more complex challenge and get practice organizing and presenting team results. The second challenge, Chems-R-Us, serves as a three-week final project. Students perform both a classification task (which compounds are readily-biodegradable or not), and feature selection task. Students work in teams as if they are consultants hired by a hypothetical company Chems-R-Us. The lab times are spent coaching the student teams instead of traditional lectures. As a team, they are required to explore a diverse set of classification and feature selection approaches and formally present their final results and recommendations in a mini-conference.

The use of the two competitions based on real-world problem has enabled the class to incorporate compelling project-based learning experiences to a large class of 60 to 90 students. We have been reusing copies of the same competitions for over five years, but students come up with new approaches and outcomes every year. Team work and creative exploration is enhanced by having two separate tasks in one challenge and by picked the winning team for each task by the team median score. The gamification aspect encourages the many students to deeply engage with the problems and try many machine learning approaches. We vary the usage of the competitions each semester slightly by asking questions that cause them to go deeper into the analysis. Students are also asked to prepare a final report which describes the methods and results as well as their own creative analyses and visualizations to address these questions. We find the students are well prepared for the following course-based undergraduate research experiences described in the next case study 8.

Another type of hackathon taking place around the clock for 24 hours is the Data Science Game, a **world-wide competition organized by students and for students**. One of our students (Benjamin Donnot) was in the organizing team, and one of us (Isabelle Guyon) was a coach and judge. This competition is open to students from first year of master up until last year of PhD. It aims both at promoting machine learning and evaluating the level of students from the best data science programs worldwide. The event took place yearly from 2015 to 2018. A first pre-selection round is run online. The best teams are then invited on-site in a castle to compete non-stop for 24 hours. Mentors from both academia and industry are invited to coach the students. Many sponsors provide prizes and travel awards. For example, the first edition gathered 143 teams from more than 50 different prestigious universities (such as Stanford, National University of Singapore, Columbia University, Cambridge University, Franch Ecole Polytechnique, and Moscow State University), from 28 different countries from 5 different continents. The level was very high as some top ranked kagglers were participating to the event. It has been supported by various sponsors such as: Google, Microsoft, AXA, CapGemini but also some other institutions such as ChaLearn or Etalab.

Hackathons also lend themselves to be organized **in conjunction with beginner tutorials at machine learning conferences**. Université Paris-Saclay co-organized a hackathon on Malaria microscopy analysis in conjunction with Data Science Africa 2019, which attracted 254 participants. One of our students (Herilalaina Rakotoarison) made a one hour presentation: 5 min (Codalab presentation), 15 min (how to create a challenge) and 40 min (mini hackathon). See slides. The 40 min hackathon helped the students learn to run the Jupyter notebook until the end. The students could then continue making entries until the end of the conference. One difficulty was that the internet connection was not very good. The winners received support to go to the next conference. All the participants were then encouraged to enter a more complex version of the challenge.

Finally, ChaLearn also held a hackathon as part of a **summer school**, the Microsoft Machine Learning and Intelligence School, Saint Petersburg, Russia, July 29 - August 5, 2015. One of our students (Arthur Pesah) was part of the organizing team. The school was sponsored by Microsoft Research and Yandex and is organized in cooperation with Lomonosov Moscow State University (MSU). It offered advanced undergraduates, PhD students, and young scientists and developers an opportunity to learn about the latest research in machine learning, intelligence and data science from top scientists. We offered a simplified version of the AutoML challenge, which helped students enter that competitions. Prizes were awarded, not only to the winners, but also to the team, which made the best presentation. The students were encouraged to continue on working on the larger international competition.

8 Case study 3: Course-based undergraduate research based on high-stake competition

Competitions can be a very effective way for educators to provide Course-Based Undergraduate Research Experiences (CUREs). CUREs are a pedagogical approach in which students engage in original research with unknown outcomes as part of a regular course. They offer a more inclusive entry point into research for undergraduates with proven benefits for students outcomes Bangera and Brownell (2014).

Students of Kristin Bennett at RPI participated in a high stake competition: the AHRQ Visualization of Community-Level Social Determinants of Health Challenge as research projects in two CUREs. This Challenge invited participants to develop new online tools to present social determinants of health data. The goal was ultimately to improve population health outcomes, and drive savings. The Challenge was structured in two phases. In Phase 1, which launched in March 2019, participants submitted concept abstracts and prototype designs of data visualization methods. For Phase 1, 12 semifinalists received $10,000 each based on the merits of their proposals and moved to Phase 2. The RPI students qualified. In Phase 2, semifinalists developed proofs-of-concept to be judged by the expert panel. One grand prize winner from Phase 2 won $50,000; second place was $35,000; third place won $15,000. The RPI students won third place, with Mortality Minder Bhanot et al. (2024). Mortality Minder explores mortality trends for midlife adults ages 25–64 across the United States from 2000-2017. Users can identify social and economic factors associated with increased mortality trends at the county level for the Nation and individual States. Visualizations demonstrate determinants and their impact on mortality trends.

A team of 23 RPI students developed the MortalityMinder datasets and analytics for the competition in the Health Analytics Challenges Lab course in the summer 2019 semester. A team of 22 students prepared the final Mortality Minder entry in Fall 2019. They designed Mortality Minder as an interactive, web-based dashboard that enables healthcare researchers, providers, payers, and policy makers to gain actionable insights into how, where, and why midlife mortality rates are rising in the United States. Students were advised by health industry professionals from United Health Foundation, Continuum Health, CDPHP, and NYSDOH as well as RPI professors and scientists. Students presented and interacted with advisers many times over the summer; group projects representing aspects of the Mortality Minder project culminated with poster presentations during a term-ending Mini Conference with guests invited. The student targeted a variety of issues ranging from data transformation, analytics, and interactive visualization through system design, user interface design, usability studies, and user documentation. The large number of students involved in the Mortality Minder project and the “production" nature of the coding effort was a unique challenge for the research advisers but also enhanced the student experience. The live Mortality Minder application is publicly available and the code public code repository is open-sourced.

We briefly summarize the work of the students. Mortality Minder illustrates midlife mortality rate increases reported in Woolf and Schoomaker (2019), while providing greater insight into community-level variations and their associated factors to help determine remedies. Using authoritative data from the CDC and other sources, Mortality Minder is designed to help health policy decision makers in the public and private sectors identify and address unmet healthcare needs, healthcare costs, and healthcare utilization. Innovative analysis divides counties into risk groups for visualization and correlation analysis using K-Means clustering and Kendall correlation. For each selected State and Cause of Death, Mortality Minder dynamically creates three analysis and visualization infographics, presented as pages in the app: “National View" reveals midlife mortality rates through time and compares state and national trends; “State View" categorizes counties into risk groups based on their midlife mortality rates over time. The app determines correlations of factors to risk groups and visualizes the most significant protective and destructive factors; “Factor View" enables users to explore individual factors including their relation to the selected cause at a county level for each state and the distribution of those factors within each state. Mortality Minder also allows users to perform a nationwide analysis.

9 Case study 4: University student organized challenges

We present an example of challenge organization class, which ended up as an accepted NeurIPS’22 competition: Cross-domain MetaDL Carrión-Ojeda et al. (2022).

As previously mentioned, we (Isabelle Guyon and Adrien Pavao) taught a class on challenge organization at the master level at Université Paris-Saclay (Section 4 and 5). The class, Creation of AI Challenges, has the objective of learning to create mini student challenges, which are then solved by other students, see a list of past challenges organized by students, as part of their class requirements. To raise awareness on issues related to ethics, privacy, and fairness, this class is taught in conjunction with a class on Fairness in Artificial Intelligence.

In 2021/2022, the fairness class was new, and it attracted a lot of interest. Many students chose to address problems of fairness and/or bias in their challenge. Due to the sensitive nature of data involving human subjects, some groups decided to study bias with data involving object recognition data, as a metaphor for societal bias. For example, using the background of an image as adjunct information to recognize an object can be used as a metaphor for using the dressing style of a person to evaluate their technical skills.

The challenge class is spread over 7 weeks in January and February, with 3 hours of class per week, including 1 hour of lecture, 1/2 hour of practical tips, and 1.5 hour of practical work. The students are also expected to work 1/2 day on their own each week. The master program also includes several project and internship requirements: one 60 hour academic project to be carried out in a research lab at the university, and one two to four month internship, either in a lab or in industry. Several students chose to do either or both of their project and internship on a subject related to challenge organization. This allowed us to involve them in the organization of an international challenge, leveraging the effort they put into their challenge class work.

The student work was incorporated in the creation of a large meta-dataset, called Meta-Album, consisting of 40 re-purposed public datasets, obtained from various source. All problems are image classification problems, where images are reduced to 128×128 pixels. Meta-Album datasets were used in the Cross-domain MetaDL challenge, co-organized by one of the students (Dustin Carrión), and accepted in April 2022 in the NeurIPS'22 conference competition program (after a very selective review process).

In November and December, five students decided to work on creating datasets in Meta-Album format, as their project, knowing that such datasets would be used in their student challenge, and would possibly be incorporated in Meta-Album and the Cross-domain MetaDL challenge. They were tutored by a second year master student (Ihsan Ullah), who took the challenge class the year before, and was the leader of the Meta-Album effort. The students' work included to hunt for suitable datasets (image datasets with sufficient resolution to be reduced to 128×128 pixels, with at least 20 classes and at least 40 examples for class for these classes, and belonging to a certain chosen domain). Then they has to format the data, document the datasets, and run baseline methods. Part of their work was also to scrutinize data to identify possible biases. For example, some datasets included images that were extracted from videos, hence were not independent of one another. Other spurious dependencies between images included images cut out from the same mother image (satellite picture of microscope slice). This such images would have correlated color spectra (due to illumination or staining).

In January and February, the students created their student challenges, which are posted on this page. All challenge protocols were in the AutoML setting: code was submitted and blindly evaluated for training and testing on an unknown dataset (a different one in each phase). Sample data and a starting kit were provided to help participants make an entry. One of the student challenges, TrustAI, was not on image classification. They addressed the problem of bias in machine learning models against groups in society. Inspired by the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) biased software, this challenge deals with a classification problem that is adjudicating on the suspect based on criminal activities. The goal is to avoid relying on protected variables such as age, gender, or race. The other two challenges were on image classification. The PACHAMAMA challenge proposed to classify living species. To purpose was to help quantifying biodiversity and monitor population changes, for a better conservation of living organisms. Problems of bias, particularly due the image background, were part of the challenge. The PANACEA challenge proposed a histology classification problem. The goal was to help histologist and histopathologist make better diagnoses, using images of microscope slices. This problem is important to automate the process of medical analyses and help study problems of bias due to differences in staining, sample contamination, lighting, etc.

In March and April, other students chose to solve these challenges as their projects, under the guidance of Adrien Pavão.

Simultaneously, three students chose to combine their project and internship into a 6-month internship to continue working on the preparation of international challenges using Meta-Album Ullah et al. (2022). Dustin Carrión joined the team preparing the Cross-domain MetaDL challenge, took leadership, and decided to submit a proposal to NeurIPS'22 (which ended up being accepted). He went on to write a paper on the design of the challenge and baseline results, which was accepted to a meta-learning workshop at ECML/PKDD 2022 Carrión-Ojeda et al. (2022). The paper describes the design and baseline results for the challenge. Meta-learning aims to leverage experience gained from previous tasks to solve new tasks efficiently (i.e. with better performance, little training data and/or modest computational resources). While previous challenges in the series focused on *within-domain* few-shot learning problems, with the aim of learning efficiently *N-way k-shot* tasks (i.e. N class classification problems with k training examples), this competition challenges the participants to solve “any-way” and “any-shot” problems drawn from various domains (healthcare, ecology, biology, manufacturing, and others) chosen for their humanitarian and societal impact. It is based on a subset of Meta-Album Ullah et al. (2022), a meta-dataset of 40 image classification datasets from 10 domains, from which we carve out tasks with any number of “ways” (within the range 2-20) and any number of “shots” (within the range 1-20). The competition is with code submission, fully blind-tested on the CodaLab challenge platform. The code of the winners will be open-sourced, enabling to deploy automated machine learning solutions for few-shot image classification across several domains.

The two other students (Gabriel Lauzzana and Romain Mussard) set on working on the preparation of a challenge on bias. Their design was submitted for publication in a junior conference Lauzzana et al. (2022). The focused on datasets, not involving human subjects, but plagued with various kinds of bias, the origin of which is not always known, and which may include confounding bias and sampling bias. They propsed to unravel causes of bias, followed by rigorous manual data curation. With the advent of fully automated machine learning (AutoML), one may wonder whether creating bias-robust learning machines is possible, to reduce the need for data curation and the possible risk of introducing further biases. In this context, they designed a bias-aware AutoML challenge, based on image classification tasks. We presented in the paper the challenge design, data preparation, and baseline methods. For reproducibility their code is provided.

In conclusion, the student competition design effort was very fruitful and resulted in the preparation of several international competitions.

10 Case study 5: Hackathons in industry - Continuous learning and upskilling

Hackathons represent an exciting way in industry to learn about newer technologies being used in the enterprise and use them to solve customer scenarios or apply acquired knowledge to new domains, or yet to remain relevant, be at the forefront of change and sometimes help advance one’s career.

As early as 2010 when ML was mature enough to being used in industry, hackathons helped bring experts and novices in industry to together learn and solve problems. Hackathons can take the form of:

- A free form track where people could work on any problem of their choice as long as they were articulate about the customer problem/scenario and that the proposed solution used AI/ML; the deliverable had to be a prototype or demo with the winners determined by a jury.
- A contest track where business groups would pitch problems to be solved and would selected the “winner” with specific goals such as ’achieving highest percentage accuracy using any tool; with the “Winners” being determined by an automated script reflected on a leaderboard.

Hackathons are today at the center stage of the culture of innovation needed to transform industries, businesses or organizations to be empowered by AI. Hackathons are becoming more than a one off tool to learn some new technologies; they are part of a larger framework within an innovation lifecycle from explorations to learning to validating ideas and building prototypes ready for scale (see `https://www.microsoft.com/en-us/garage/`).

Google DeepMind also organizes yearly hackathons, which used tom be called “DeepMind extravaganza" (now GDM-create). It is a two-week annual event focused on cross-team collaboration, learning, and connection. During that time, team members have the opportunity to engage in exploratory projects that they propose or are proposed by others. It fosters a continuous learning culture within DeepMind.

11 Conclusion

In this chapter, we have shown that competitions and challenges offer a powerful pedagogical approach across all stages of learning and professional development. From the earliest levels of education, where activities often take the form of simple, juried projects, to university courses leveraging automated scoring on complex research problems, the competition format promotes engagement, self-directed learning, and creative problem-solving. Additionally, we have seen how competitions can serve as effective hands-on laboratories for teaching experimental design, data analysis, and application of cutting-edge techniques.

Beyond academia, competitions and challenges also support continuous learning in industry. They facilitate the learning of new methods, tools, and technologies, helping professionals maintain their relevance in a fast-evolving marketplace. By encouraging learners to embrace hands-on problem-solving and exploration, these activities help bridge the gap between theory and practice, and they can be adapted to a wide range of domains and skill levels.

In sum, competitions and challenges provide an adaptable and practical way to motivate learners. As educational and professional environments continue to change, their value as both a teaching strategy and a continuous learning mechanism will likely only increase.

Acknowledgements

We thank Tara Chklovski for her invaluable contributions.

References

Risqa D Anazifa and Djukri Djukri. Project-based learning and problem-based learning: Are they effective to improve student's thinking skills? *Jurnal Pendidikan IPA Indonesia*, 6(2):346–355, 2017.

Albert Bandura. *Self-Efficacy: The Exercise of Control*. W. H. Freeman, 1997.

Gita Bangera and Sara E Brownell. Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, 13(4):602–606, 2014.

Kristin P Bennett, John S Erickson, Amy Svirsky, and Josephine C Seddon. A mathematics pipeline to student success in data analytics through course-based undergraduate research. *The Mathematics Enthusiast*, 19(3):730–750, 2022.

Karan Bhanot, John S Erickson, and Kristin P Bennett. Mortalityminder: Visualization and ai interpretations of social determinants of premature mortality in the united states. *Information*, 15 (5):254, 2024.

Paolo Calafiura, Steven Farrell, Heather M. Gray, Jean-Roch Vlimant, Vincenzo Innocente, Andreas Salzburger, Sabrina Amrouche, Tobias Golling, Moritz Kiehn, Victor Estrade, Cécile Germain, Isabelle Guyon, Ed Moyse, David Rousseau, Yetkin Yilmaz, Vladimir Vava Gligorov, Mikhail Hushchyn, and Andrey Ustyuzhanin. Trackml: A high energy physics particle tracking challenge. In *14th IEEE International Conference on e-Science, e-Science 2018, Amsterdam, The Netherlands, October 29 - November 1, 2018*, page 344. IEEE Computer Society, 2018. doi: 10.1109/eScience.2018.00088. URL https://doi.org/10.1109/eScience.2018.00088.

Dustin Carrión-Ojeda, Hong Chen, Adrian El Baz, Sergio Escalera, Chaoyu Guan, Isabelle Guyon, Ihsan Ullah, Xin Wang, and Wenwu Zhu. Neurips'22 cross-domain metadl competition: Design and baseline results. *CoRR*, abs/2208.14686, 2022. doi: 10.48550/arXiv.2208.14686. URL https://doi.org/10.48550/arXiv.2208.14686.

Dustin Carrión-Ojeda, Hong Chen, Adrian El Baz, Sergio Escalera, Chaoyu Guan, Isabelle Guyon, Ihsan Ullah, Xin Wang, and Wenwu Zhu. In *Meta-Knowledge Transfer/Communication in Different Systems*, ECML/PKDD 2022 workshop, 2022.

Tara Chklovski, Richard Jung, Rebecca Anderson, and Kathryn Young. Comparing 2 years of empowering families to solve real-world problems with ai. *KI-Künstliche Intelligenz*, 35(2):207–219, 2021.

Anthony Goldbloom and Ben Hamner. Kaggle. 2010. URL https://www.kaggle.com/docs/competitions.

Suzanne Hidi and K. Ann Renninger. The four-phase model of interest development. *Educational Psychologist*, 41:111 – 127, 2006.

Robyn Hromek and Sue Roffey. Promoting social and emotional learning with games: "it's fun and we learn things". *Simulation & gaming*, 40(5):626–644, 2009.

Dirk Ifenthaler, Deniz Eseryel, and Xun Ge. Assessment for game-based learning. In *Assessment in game-based learning*, pages 1–8. Springer, 2012.

Balázs Kégl, Alexandre Boucaud, Mehdi Cherti, Akin Osman Kazakci, Alexandre Gramfort, Guillaume M Lemaitre, Joris Van den Bossche, Djalel Benbouzid, and Camille Marini. The RAMP framework: from reproducibility to transparency in the design and optimization of scientific workflows. International Conference On Machine Learning, July 2018. URL https://hal.archives-ouvertes.fr/hal-02072341. Poster.

Gabriel Lauzzana, Romain Mussard, Ihsan Ullah, and Isabelle Guyon. Design of a bias-aware automl challenge. In *JDSE*, In review, 2022.

Zhengying Liu, Adrien Pavao, Zhen Xu, Sergio Escalera, Fabio Ferreira, Isabelle Guyon, Sirui Hong, Frank Hutter, Rongrong Ji, Julio C. S. Jacques Junior, Ge Li, Marius Lindauer, Zhipeng Luo, Meysam Madadi, Thomas Nierhoff, Kangning Niu, Chunguang Pan, Danny Stoll, Sebastien Treguer, Jin Wang, Peng Wang, Chenglin Wu, Youcheng Xiong, Arbër Zela, and Yang Zhang. Winning solutions and post-challenge analyses of the ChaLearn AutoDL challenge 2019. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, page 17, 2020.

Antoine Marot, Benjamin Donnot, Camilo Romero, Luca Veyrin-Forrer, Marvin Lerousseau, Balthazar Donon, and Isabelle Guyon. Learning to run a power network challenge for training topology controllers. *CoRR*, abs/1912.04211, 2019. URL http://arxiv.org/abs/1912.04211.

Antoine Marot, Benjamin Donnot, Gabriel Dulac-Arnold, Adrian Kelly, Aïdan O'Sullivan, Jan Viebahn, Mariette Awad, Isabelle Guyon, Patrick Panciatici, and Camilo Romero. Learning to run a power network challenge: a retrospective analysis. *CoRR*, abs/2103.03104, 2021. URL https://arxiv.org/abs/2103.03104.

Hoa P. Nguyen. How to use gameplay to enhance classroom learning. *Edutopia*, 2021. URL https://www.edutopia.org/article/how-use-gameplay-enhance-classroom-learning.

Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.

Adrien Pavao, Diviyan Kalainathan, Lisheng Sun-Hosoya, Kristen Bennett, and Isabelle Guyon. Design and analysis of experiments: A challenge approach in teaching. In *NeurIPS 2019-33th Annual Conference on Neural Information Processing Systems*, 2019.

Adrien Pavao, Isabelle Guyon, Anne-Catherine Letournel, Xavier Baró, Hugo Escalante, Sergio Escalera, Tyler Thomas, and Zhen Xu. Codalab competitions: An open source platform to organize scientific challenges. *Technical report*, 2022. URL https://hal.inria.fr/hal-03629462v1.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Jan L Plass, Bruce D Homer, and Charles K Kinzer. Foundations of game-based learning. *Educational psychologist*, 50(4):258–283, 2015.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL https://www.R-project.org/.

Ihsan Ullah, Dustin Carrion, Sergio Escalera, Isabelle Guyon, Mike Huisman, Felix Mohr, Jan N. van Rijn, Haozhe Sun, Joaquin Vanschoren, and Phan Anh Vu. Meta-album: Multi-domain meta-dataset for few-shot image classification. In *Submitted to: Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, 2022. URL https://meta-album.github.io/.

Steven H Woolf and Heidi Schoomaker. Life expectancy and mortality rates in the united states, 1959-2017. *Jama*, 322(20):1996–2016, 2019.

W-H Wu, H-C Hsiao, P-L Wu, C-H Lin, and S-H Huang. Investigating the learning-theory foundations of game-based learning: a meta-analysis. *Journal of Computer Assisted Learning*, 28(3): 265–279, 2012.

Andrew Yale, Saloni Dash, Ritik Dutta, Isabelle Guyon, Adrien Pavao, and Kristin P Bennett. Generation and evaluation of privacy preserving synthetic health data. *Neurocomputing*, 416:244–255, 2020.