

# **CiML 2015**

## **Challenges (competitions) in Machine Learning: Open Innovation and Coopetitions**

**NIPS 2015 Workshop**

**Organisers:** Isabelle Guyon, ChaLearn; Ben Hamner, Kaggle;  
Balázs Kégl, CNRS-Paris; Evelyne Viegas, Microsoft

# Competitions

## Some milestones



Longitude Prize , 1714



DARPA Grand Challenge 2005



DARPA Network Challenge 2009

## Bring research to industry via open innovation

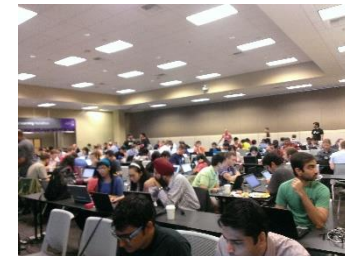


Netflix Recommendation



GoldCorp Challenge

## Upskilling in industry



Hackathon

# The value of Challenges

- Means of directing research, advancing the state-of-the-art or venturing in completely new domains
- Playful nature naturally attracts students, making challenge a great teaching resource, and information workers wanting to upskill
- Efficient and cost-effective ways to quickly bring to industry solutions that may have been confined to research
- Participants, undergraduates to retirees, work in a rewarding environment where they can learn, perform research, demonstrate excellence
- Challenges have become streamline in Machine Learning projects, it has become increasingly important to regularly bring together workshop organizers, platform providers, and participants to discuss
  - best practices in challenge organization
  - new methods and application opportunities to design high impact challenges

# CiML'15 – informed from CiML'14

**Open innovation** – data made available and contestants must both formalize and solve a problem, leaving more freedom to creativity, while giving more difficulty to the organizers to objectively assess the results

**Coopetitions** – encourage both collaboration and competition, and make possible the contributions of many towards solving an overall problem; this poses to the organizers the problem of rewarding partial contributions

**Platforms** – platforms and protocols permitting code submission, not just result submission, allowing fairer standardized comparisons in terms of hardware utilization and easier reproducibility and "worksheets" or "scripts" facilitating code sharing

**Sharing, dissemination, and recognition** – Facilitate sharing resources, including data, means of data collection and annotation, challenge announcements, best practices, challenge templates, publication channels, etc.; creation of awards to recognize academic services rendered by the various actors of challenge organization

# Discussions – Voting Rules

- Stickers
  - Each person should have 3 stickers
  - Use them on any of the poster (up to 3)
- Post-its
  - Add your top burning questions to any of the poster
  - Any new category, add them to the Open discussion poster

# CiML'15 Agenda

## Morning session (9:00 am-12:00 pm)

9:00 - **Welcome and introduction.** E. Viegas

9:10 - **Challenges in Medical Image Analysis: Comparison, Competition, Collaboration,** B. van Ginneken

**9:50 - Break Add your VOTE for Discussion session**

10:20 - **Techniques and Technologies for Efficient and Realistic Benchmarks: Examples from the MediaEval Multimedia Benchmark and CLEF NewsREEL,** M. Larson

11:00 - **Discussion: Open Innovation,** B. Kegl and B. Hamner moderators

12:00 – 14:30 Break-out session on [AutoML challenge](#). Free lunch for the participants

**AutoML challenge.** I. Guyon -- Announcement of the [new GPU track](#).

**Automated Machine Learning: Successes & Challenges.** F. Hutter. Team aaad\_freiburg. First in AutoML1, second in AutoML2

**Sensible allocation of computation for ensemble construction.** J. Lloyd. (jrl44/backstreet.bayes). First AutoML2, second AutoML1

**Scalable ensemble learning with stochastic feature boosting.** E. Tuv. Team ideal.intel.analytics. First place Final0, second Final1

**14:30 - Break (30 min) - Add your VOTE for Discussion session**

## Afternoon session (15:00-18:30)

15:00 - **Lessons Learned from the PASCAL VOC Challenges, Improving the Data Analytics Process,** C. Williams

15:40 - **Discussion: Coopetitions,** E. Viegas and I. Guyon moderator

16:40 - Break (20 min)

17:00 - **Academic Torrents: Scalable Data Distribution,** H. Z. Lo and J. Paul Cohen

17:30 - **Open discussion,** M. Sebag, modelator

18:20 - Wrap up