



NIPS 2018 Challenge

The 3rd AutoML Challenge: AutoML for Lifelong Machine Learning

(Provided and Sponsored by 4Paradigm, ChaLearn, Microsoft and Acadia University)

[Overview](#)
[Data](#)
[Prizes](#)
[Rules](#)
[Timeline](#)
[About](#)

Overview

The competition has been launched at CodaLab, please follow the link to participate: <https://competitions.codalab.org/competitions/19836>

In many real-world machine learning applications, AutoML is strongly needed due to the limited machine learning expertise of developers. Moreover, batches of data in many real-world applications may be arriving daily, weekly, monthly, or yearly, for instance, and the data distribution are changing relatively slowly over time. This presents a continuous learning, or Lifelong Machine Learning challenge for an AutoML system. Typical learning problems of this kind include customer relationship management, on-line advertising, recommendation, sentiment analysis, fraud detection, spam filtering, transportation monitoring, econometrics, patient monitoring, climate monitoring, manufacturing and so on. In this competition, which we are calling **AutoML for Lifelong Machine Learning**, large scale datasets collected from some of these real-world applications will be used. Compared with previous AutoML competitions (<http://automl.chalearn.org/>), the focus of this competition is on **drifting concepts**, getting away from the simpler i.i.d. cases. Participants are invited to design a computer program capable of autonomously (without any human intervention) developing predictive models that are trained and evaluated in a lifelong machine learning setting.

Although the scenario is fairly standard, this challenge introduces the following difficulties:

- **Algorithm scalability.** We will provide datasets that are 10–100 times larger than in previous challenges we organized.
- **Varied feature types.** Varied feature types will be included (continuous, binary, ordinal, categorical, multi-value categorical, temporal). Categorical variables with a large number of values following a power law will be included.
- **Concept drift.** The data distribution is slowly changing over time.
- **Lifelong setting.** All datasets included in this competition are chronologically splitted into 10 batches, meaning that instance batches in all datasets are chronologically ordered (note that instances in one batch are not guaranteed to be chronologically ordered). The algorithms will be tested for their capability of adapting to changes in data distribution by exposing them to successive test batches chronologically ordered. After testing, the labels will be revealed to the learning machines and incorporated in the training data.

There're two phases of the competition:

The **Feedback phase** is a phase with **code submission**, participants can practice on 5 datasets that are of similar nature as the datasets of the second phase. Participants can make a limited number of submissions. Participants can download the labeled training data and the unlabeled test set. So participants can prepare their code submission at home and submit it later. The LAST code submission will be forwarded to the next phase for final testing.

The **AutoML phase** is the blind test phase with **no submission**. The last submission of the previous phase is blind tested on 5 new datasets. Participant's code will be trained and tested automatically, without human intervention. The final score will be evaluated by the result of the blind testing.

Mirror Sites for Public Dataset:France: https://newcodalab.lri.fr/automl3/AutoML3_input_data_1.zipMexico: https://ccc.inaoep.mx/~hugojair/AutoML3_input_data_1.zipChina: <https://pan.baidu.com/s/1Bm4gAGwtCXLiM72kuwKDjg> Password: **wv8a**

For **WeChat** Users, the official WeChat Group for this competition: (Because the group is over 100 members, new members should be invited to group.)

**Data****Data Format**

For each instance, we have following 4 types of features splitted by blank in our instance files, i.e.,

- *Categorical* Feature: an integer describing which category the instance belongs to.
- *Numerical* Feature: a real value.
- *Multi-value Categorical* Feature: a set of integers, splitted by comma. The size of the set is not fixed and can be different for each instance. For example, topics of an article, words in a title, items bought by a user and so on.
- *Time* Feature: an integer describing time information.

Note: Categorical/Multi-value Categorical features with a large number of values following a power law might be included.

Feedback Phase Public Datasets

5 public datasets are released, including their first 5 batches, another 5 batches are kept private.

| DATASET | Budget | #Cat | #Num | #MVC | #Time | #Feature | #Instance |
|---------|--------|------|------|------|-------|----------|--------------|
| A | 3600 | 51 | 23 | 6 | 2 | 82 | ~10 Million |
| B | 600 | 17 | 7 | 1 | 0 | 25 | ~1.9 Million |
| C | 1200 | 44 | 20 | 9 | 6 | 79 | ~2 Million |
| D | 600 | 17 | 54 | 1 | 4 | 76 | ~1.5 Million |
| E | 1800 | 25 | 6 | 1 | 2 | 34 | ~17 Million |

Budget=Time budget(seconds). #Cat=Number of categorical features. #Num=Number of numerical features. #MVC=Number of multi-value categorical features. #Time=Number of time features. #Feature=Total number of features. #Instance=Total number of instances for all 10 batches.

AutoML Phase Private Datasets

5 private datasets to be released.

Basic Tips

Here, we provide some basic tips for dealing with large datasets:

- Subsampling/multi-fidelity AutoML approaches might be needed for these datasets.
- Incremental learning might be needed for these datasets.

Some basic tips for handling difficult features:

- One-hot encoding for Categorical features.
- Hashing tricks might be used for Categorical and Multi-value Categorical features.
- Normalization tricks for Numerical features.
- If the number of elements in a Categorical/Multi-value Categorical feature is too large, instead of hashing tricks, one might compute moving frequencies of these elements and always keep top ones.
- For Time features, one might minus a fixed value from the original feature. When multiple Time features present, one might construct new features based on the difference of these Time features.
- For missing features, i.e., NaN, one might set it as a default value or replace it with another valid value indicating it is missing.

Note: The competition focuses on automatically combating with the drifting concepts. The processing of features might need to be adaptive over time. Automatic feature generation and selection methods or deep learning approaches might be important for these datasets.

Prizes

AutoML 1st Place \$10000(+ *Travel Grant*)

2nd Place \$3000(+ *Travel Grant*)

3rd Place \$2000(+ *Travel Grant*)

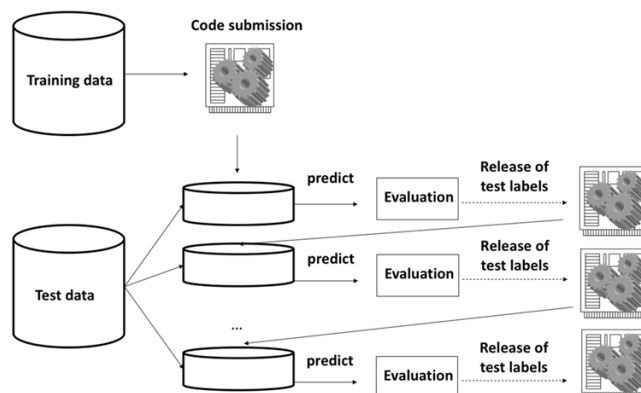
* Champions will provide travel grant to attend the conference and workshop.

Rules

Evaluation

The evaluation scheme is depicted in image below. Recall this is a competition receiving **code submissions**. Participants must prepare an AutoML program that will be uploaded to the challenge platform. The code will be executed in computer workers, autonomously; and allowed to run for maximum amount of time. Code exceeding this time will be penalized with setting the dataset's AUC as 0. Different from previous challenges, this competition we will evaluate the Lifelong learning capabilities of AutoML solutions, hence an appropriate protocol has been designed.

The datasets are chronologically split into 10 batches, each batch will represent a stage of the lifelong evaluation scenario. Code submitted by participants will use the first batch to generate a model, which will then be used to predict labels for the first test batch(i.e., the second batch). performance on this test batch will be recorded. After this, the labels of the first test batch will be made available to the computer program. The computer program may use such labels to improve its initial model and make predictions for the subsequent test batch. The process will continue until all of the test batches have been evaluated. We call this **1 / 9 split evaluation**, meaning that first batch for initial training, all successive 9 batches for evaluation.



| STEP# | TRAINING DATA | TEST DATA |
|-------|---------------------------------------|-------------------|
| 1 | LABELED BATCH_0 | UNLABELED BATCH_1 |
| 2 | LABELED (BATCH_0 + BATCH_1) | UNLABELED BATCH_2 |
| 3 | LABELED (BATCH_0 + BATCH_1 + BATCH_2) | UNLABELED BATCH_3 |
| ... | ... | ... |
| 9 | EVERYTHING LABELED UP TO BATCH_8 | UNLABELED BATCH_9 |

- For each dataset, the first batch will be released as initial training set, the evaluation will be operated on all the rest 9 batches.
- For each batch of each dataset, the evaluation will consist in computing the area under the ROC curve (AUC).
- For each dataset, we will take the average of the AUC ranks over all the successive 9 batches of the dataset. A ranking will be performed according to this metric.
- For the final score, we will use the average rank over all datasets.
- There will be a time budget for each dataset. Code exceeding the maximum execution time will be aborted and assigned an AUC=0.

Note:

- The **1 / 9 split evaluation** will be used in both Feedback and AutoML phases.
- But all datasets in Feedback phase consist of 5 released batches + 5 private batches. Although 5 batches are released at Feedback Phase, the evaluation will be calculated at all the 9 batches except the first batch. This is the same as the AutoML phase.

The rationale:

- During Feedback phase, participants are provided with 5 released batches so that they can develop their method with at home, but when they submit their code, they are evaluated identically in both phases.
- During AutoML phase, we need as many test batches as possible to see the significance of the adaptation.
- In Feedback phase, the test on the first 4 batches will be biased. We will show detailed results on all batches, so it will be evident if there is bias or not.

Terms & Conditions

- The competition will be run in the CodaLab competition platform
- The competition is open for all interested researchers, specialists and students. Members of the Contest Organizing Committee cannot participate.
- Participants may submit solutions as teams made up of one or more persons.
- Each team needs to designate a leader responsible for communication with the Organizers.
- One person can only be part of one team.
- A winner of the competition is chosen on the basis of the final evaluation results. In the case of draws in the evaluation scores, time of the submission will be taken into account.
- Each team is obligated to provide a short report (fact sheet) describing their final solution.
- By enrolling to this competition you grant the organizers rights to process your submissions for the purpose of evaluation and post-competition research.

Dissemination

- Top ranked participants will be invited to attend a workshop collocated with NIPS 2018 to describe their methods and findings. Winners of prize are expected to attend.
- The challenge is part of the competition program of the NIPS 2018 conference. Organizers are making arrangements for the possible publication of a book chapter or article written jointly by organizers and the participants with the best solutions.

Timeline[Tentative]

2nd March, 2018: Preparation of the competition starts.

30th July, 2018: Beginning of the competition, release of development data.

23rd October, 2018: End of development (Feedback) phase.

6th November, 2018: End of the competition.

13th November, 2018: Deadline for submitting the fact sheets.

20th November, 2018: Release of results, announcements of winners.

November 2018: Post challenge analyses.

December, 2018: NIPS2018. Autonomous Machine Learning for Lifelong Machine Learning Ceremony

About

Sponsors



Committee

In case of any questions please send an email to automl2018@gmail.com.

Wei-Wei Tu, 4Paradigm Inc., Beijing, China, (Coordinator, Baseline Provider, Data Provider), tuww.cn@gmail.com

Hugo Jair Escalante, INAOE (Mexico), ChaLearn (USA), (Platform Administrator, Coordinator), hugo.jair@gmail.com

Isabelle Guyon, UPSud/INRIA Univ. Paris-Saclay, France & ChaLearn, USA, (Coordinator, Platform Administrator, Advisor), guyon@clopinet.com

Daniel L. Silver, Acadia University, (Advisor), danny.silver@acadiau.ca

Evelyne Viegas, Microsoft Research, (Coordinator, Advisor), evelynev@microsoft.com

Yuqiang Chen, 4Paradigm Inc., Beijing, China, (Sponsor, Data Provider), chenyuqiang@4paradigm.com

Qiang Yang, 4Paradigm Inc., Beijing, China, (Advisor, Sponsor), qyang@cse.ust.hk

Sinno Jialin Pan, NTU Singapore, (Advisor), sinnopan@ntu.edu.sg

Quanming Yao, 4Paradigm Inc. Beijing, China, (Baseline Provider, Data Provider) yaoquanming@4paradigm.com

Mengshuo Wang, 4Paradigm Inc. Beijing, China, (Baseline Provider, Data Provider) wangmengshuo@4paradigm.com

Yuanyu Wan, 4Paradigm Inc. Beijing, China, (Baseline Provider, Data Provider) wanyuanyu@4paradigm.com

Hai Wang, 4Paradigm Inc. Beijing, China, (Baseline Provider, Data Provider) wanghai@4paradigm.com

Organization Institutes



About AutoML Challenge

Previous AutoML Challenges: [The First AutoML Challenge](#) and [The Second AutoML Challenge](#).

AutoML workshops can be found [here](#).

Microsoft research blog post on AutoML Challenge can be found [here](#).

KDD Nuggets post on AutoML Challenge can be found [here](#).

I. Guyon et al. *A Brief Review of the ChaLearn AutoML Challenge: Any-time Any-dataset Learning Without Human Intervention*. ICML W 2016. [link](#)

I. Guyon et al. *Design of the 2015 ChaLearn AutoML challenge*. IJCNN 2015. [link](#)

Springer Series on Challenges in Machine Learning. [link](#)

About 4Paradigm Inc. (Main Sponsor, Baseline Provider & Data Provider, Coordinator)

Founded in early 2015, 4Paradigm (<https://www.4paradigm.com/>) is one of the world's leading AI technology and service providers for industrial applications. 4Paradigm's flagship product – *the AI Prophet* – is an AI development platform that enables enterprises to effortlessly build their AI applications, and thereby significantly increase their operation's efficiency. Using the *AI Prophet*, a company can develop a data-driven "AI C System", which could be largely regarded as a second core system next to the traditional transaction-oriented Core Banking System (IBM Mainframe) often found in banks. Beyond this, 4Paradigm has also successfully developed more than 100 AI solutions for use in various settings such as finance, telecommunication and Internet applications. These solutions include, but are not limited to, smart pricing, real-time anti-fraud systems, precision marketing, personalized recommendation and more. And while, it is clear that 4Paradigm can completely set up a new paradigm that an organization uses its data, its scope of services does not stop there. 4Paradigm uses state-of-the-art machine learning technologies and practical experiences to bring together a team of experts ranging from scientists to architects. This team has successfully built China's largest machine learning system and the world's first commercial deep learning system. However, 4Paradigm's success does not stop there. With its core team pioneering the research of "Transfer Learning," 4Paradigm takes the lead in this area, and as a result, has drawn great attention by world's tech giants.

About ChaLearn & CodaLab (Platform Provider, Coordinator)

ChaLearn (<http://chalearn.org>) is a non-profit organization with vast experience in the organization of academic challenges. ChaLearn is interested in all aspects of challenge organization, including data gathering procedures, evaluation protocols, novel challenge scenarios (e.g., competition training for challenge organizers, challenge analytics, results dissemination and, ultimately, advancing the state-of-the-art through challenges). ChaLearn is collaborating with the organization of the NIPS 2018 data competition (AutoML Challenge 2018).

The competition will be run in the CodaLab platform (<https://competitions.codalab.org/>). CodaLab is an open-source web-based platform that enables researchers, developers, and data scientists to collaborate, with the goal of advancing research fields where machine learning and advanced computation is used. CodaLab offers several features targeting reproducible research. In the context of the AutoML Challenge 2018, CodaLab is the platform that will allow the evaluation of participants' solutions. CodaLab is administered by Université Paris-Saclay and maintained by CKcollab, LLC. This will be possible by the funding of 4Paradigm and a Microsoft Azure for Research grant.

About Microsoft Research (Server Provider, Advisor)

At Microsoft, we aim to empower every person and every organization on the planet to achieve more. We care deeply about having a global perspective and making a difference in lives and organizations in all corners of the planet. This involves playing a small part in the most fundamental of human activities: Creating tools that enable each of us along our journey to become something more. Our mission is grounded both in the world in which we live and the future we strive to create. Today, we live in a mobile-first, cloud-first world, and we aim to enable our customers to thrive in this world. <https://www.microsoft.com/en-us/research/>

About Lifelong Machine Learning and Reasoning Group (Advisor)

The Lifelong Machine Learning and Reasoning Group (<http://lmlr.acadiau.ca/>) at Acadia University undertakes research into novel machine learning algorithms and their application to synthetic and real-world problems. The lab's researchers and students specialize in developing machine learning algorithms in the area of Lifelong Machine Learning, Transfer Learning, Knowledge Consolidation, Multi-modal Learning, and Learning to Reason. The group has particular expertise in artificial neural networks and deep learning used for supervised, unsupervised, semi-supervised recurrent (time series) problems. We also apply standard machine learning methods as well as more advanced LMLR approaches to problems in areas of data analytics, adaptive systems, intelligent agents and robotics. Most notably our lab created context sensitive Multiple Task Learning (csMTL) networks in 2005 and have advanced them for use in Lifelong Machine Learning research and application. To review some of our most recent work on Multi-modal Learning using deep learning architectures please try the demo at <https://ml3cpu.acadiau.ca/>.





关于我们

- 相关新闻
- 范式荣誉
- 社会招聘
- 校园招聘

友情链接

- 范式融通
- 智能推荐系统
- 瑞金医院

第四范式

商务合作

- 商务邮箱 contact@4paradigm.com
- 联系电话 010-82780800
- 合作咨询 ai_partner@4paradigm.com

关注我们