Codalab
Open source platform to accelerate reproducible computational research

Evelyne Viegas, Isabelle Guyon, Sergio Escalera
CodaLab 2017 Community Leads

Percy Liang, Stanford

Isabelle Guyon, Paris-Saclay, ChaLearn

Evelyne Viegas, Microsoft AI & Research

Sergio Escalera, U. of Barcelona, CV Center, ChaLearn

Xavier Baró Solé, U. Oberta de Catalunya, CV Center

Eric Carmichael, CK Collab
CodaLab – Contributors

CodaLab has received important contributions from many people, and we would like to thank their efforts in making CodaLab what it is today:

What are the issues?

- Duplication of Effort / No Shared Learning
- Lack of Reproducibility / Inadequate Publication
- No Comparable Baseline / Insufficient Collaboration
- Instant Gratification! Inability to leverage work from challenges
What is needed

- **Reducing preparation time**
  - Accelerating the pace of innovation

- **Reproducibility**
  - Enabling sharing of approach, code, data

- **Benchmarking**
  - Establishing common measures

- **Coopetitions**
  - Solving challenges collaboratively
CodaLab ecosystem

- Upload Data & Algo to CodaLab
- Conduct Data-Driven Experiment
- Publish Experimental Results
- Peer Review & Verification
- Copy & Modify Experiment
- Participate in Coopetition
- Sponsor Coopetition

Coopetition

Participate in Coopetition

Sponsor Coopetition

Upload Data & Algo to CodaLab

Conduct Data-Driven Experiment

Publish Experimental Results

Copy & Modify Experiment
2013 – CodaLab 1.0, an open source platform to accelerate reproducible computational research. Accepts result submission.

2014 – Expansion to CV, Speech Accepts submission of code (test only, no training)


CodaLab 2017 and beyond

~ 480 challenges (145 public)
~ 10000 users

See [http://www.chalearn.org/tips.html](http://www.chalearn.org/tips.html) for tips on how to organize a competition

Codalab v1.5

- Lower barrier of entry with ChaLab
- Scalability & reusability
- Enables new types of competitions (e.g. Reinforcement Learning)
Open source: [https://github.com/codalab/codalab-competitions](https://github.com/codalab/codalab-competitions)
Public instance: [https://competitions.codalab.org/](https://competitions.codalab.org/)
Local computer or cloud: Azure for research grant

Users:

Code submission [http://automl.chalearn.org](http://automl.chalearn.org)

Organizers:

Competition bundles
Competition bundles

https://github.com/codalab/competition-examples
Upload bundle
TA DAH!

Iris


Organized by guyon - Current server time: Nov. 27, 2016, 9:01 p.m. UTC

Previous
Development
Nov. 19, 2016, 10:20 p.m. UTC
Current
Final
Nov. 22, 2016, 10:20 p.m. UTC
End
Competition Ends
Never

Learn the Details
Phases
Participate
Results
Public Submissions
Forums

Overview
Evaluation
Terms and Conditions

Iris

Brought to you by ChaLearn

The famous Iris classification problem of Sir Ronald Fisher

This challenge was generated using chalab.
Wizard: ChaLab (beta) [https://chalab.lri.fr](https://chalab.lri.fr)

- Dumps
- Personal compute workers
- Dockers
- Ingestion program


[DOC-URL]
# Student challenges

[https://sites.google.com/a/chalearn.org/saclay/](https://sites.google.com/a/chalearn.org/saclay/)

<table>
<thead>
<tr>
<th>Challenges (M2 students)</th>
<th>Abstract</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VISION</strong></td>
<td>Autonomous vehicles will become a common means of transportation very soon. However, obstacles remain to be overcome, in particular obstacle avoidance. This requires powerful computer vision algorithms. In this challenge you will contribute to solve the problem of recognizing animals and vehicles.</td>
<td>To illustrate this problematic, we propose to study the image source CIFAR-10 which groups entities that can interact with the vehicle environment like animals (cat, horse, dog, ...) and vehicles (bike, car, truck, ...). We preprocessed the images to you get to solve a multi-class classification problem from pre-computed features.</td>
</tr>
<tr>
<td><strong>BIOMED</strong></td>
<td>Over-prescription of opioid medicines presents a new public health problem because many people have become addicted. This challenge asks you to help predicting which doctors tend to over-prescribe such medicines.</td>
<td>The data set contains a binary classification task. The target represents, for each medical prescription whether an opioid has been prescribed or not. The features represent, amongst others, the specialty of the doctor who made the prescription and the name of the non-opioid drugs present in this prescription.</td>
</tr>
<tr>
<td><strong>FRIEND</strong></td>
<td>Predicting at which price a house will sell helps people selling their property at a fair price. This dataset contains house sale prices for King County,</td>
<td>This is a regression problem. The dataset contains 19 house features plus the price and the id columns, along with 21613 observations.</td>
</tr>
</tbody>
</table>
Easy, but not stronger than its weakest link
http://chalab.lri.fr/
**Iris**

From group: M2AIC
Organized by ChaLearn

The well-known Iris dataset from Fisher's classic paper (Fisher, 1936).

**Challenge Definition:**

1. **Data**
   Define your data.

2. **Split**
   Define how to split data.

3. **Problem**
   Define the ingestion program.
4. Metric Definition

Name of the metric function:

mse_metric

Description:

Mean-square error

Code:

```python
# Example of organizer-provided metric.
# You can just replace this code by your own.
# Make sure to indicate the name of the function that you chose as metric function
# in the file metric.txt. E.g. mse_metric, because this file may contain more
# than one function, hence you must specify the name of the function that is your metric.
import numpy as np
import scipy as sp

def mse_metric(solution, prediction):
    """Mean-square error. Works even if the target matrix has more than one column""
    mse = np.mean((solution-prediction)**2)
    return np.mean(mse)
```

Use a Public Metric:

Load one of your Metrics

No metrics yet...

Load your Metric from a file:

Browse... No file selected.

Load
Iris

From group: M2AIC
Organized by ChaLearn
The well known Iris dataset from Fisher's classic paper (Fisher, 1936).

Challenge Definition:

1. Data
   Define your data set.

2. Split
   Define how to split data.

3. Problem
   Define the ingestion program.

4. Metric
   Define the metrics.

5. Protocol
   Define the protocol.

6. Baseline
   Define the baseline.

7. Documentation
   Document your challenge.

Package & Publish
   Package & Publish your challenge.

Compile bundle
Your own compute workers

Front end

Back end

[DOC-URL]/User_Using-your-own-compute-workers
Ingestion program

[DOC-URL]/User_Building-an-Ingestion-Program-for-a-Competition
Interactive tasks (e.g. RL)

[DOC-URL]/User_Building-an-Ingestion-Program-for-a-Competition
Codalab competitions

~150 public competitions during last 4 years

- ChaLearn is a non-profit organization focusing on challenges organization in Machine Learning
- ChaLearn Looking at People (ChaLearn LAP) is a branch of ChaLearn focusing on Human Analysis.

http://chalearn.org

http://chalearnlap.cvc.uab.es/
Chalearn LAP

Associated special issues:
• Journal of Machine Learning Research: SI-Gesture recognition 2012-2014,
• International Journal of Computer Vision: SI-Looking at People 2014-2016,
• IEEE Trans. on Affective Computing: SI-Personality Analysis
• IEEE Trans. on PAMI: SI-The Computational Face 2017
Chalearn LAP – Montalbano multimodal gesture recognition

- ECCV’14 – Pose Recovery, Action and Gesture Recognition
- ICMI’13 – Multimodal gesture recognition
- ICPR’12 – One-shot gesture recognition
- CVPR’12 – One-shot gesture recognition
- CVPR’11 – Gesture Recognition

- +14,000 gestures,
- 20 Italian sign gesture categories,
- RGB, depth, audio, subject mask and skeleton information,
- ~1.7M manually labeled frames

• Sergio Escalera, Xavier Baro, Jordi Gonzalez, Miguel A. Bautista, Meysam Madadi, Miguel Reyes, Victor Ponce, Hugo J. Escalante, Jamie Shotton, Isabelle Guyon, ChaLearn Looking at People Challenge 2014: Dataset and Results, ChaLearn Looking at People, European Conference on Computer Vision, 2014.
Chalearn LAP – Action, interaction and poses

✓ ECCV’14 – Pose Recovery, Action and Gesture Recognition

- 7 action categories,
- 4 interaction categories,
- RGB, pose and pixel labeling precision information,
- ~8K labeled frames,
- 14 limb labels.

*Sergio Escalera, Xavier Baro, Jordi Gonzalez, Miguel A. Bautista, Meysam Madadi, Miguel Reyes, Víctor Ponce, Hugo J. Escalante, Jamie Shotton, Isabelle Guyon, ChaLearn Looking at People Challenge 2014: Dataset and Results, ChaLearn Looking at People, European Conference on Computer Vision, 2014.

Chalearn LAP – Cultural event recognition

- ICCV’15 - Apparent Age Estimation and Cultural Event Recognition
- CVPR’15 - Pose Recovery, Action Recognition, and Cultural Event Recognition

- 100 cultural events,
- +27 countries around the world,
- 28K labeled RGB data.

*Sergio Escalera, Junior Fabian, Pablo Pardo, Xavier Baro, Jordi Gonzalez, Hugo Escalante, Dusan Misevic, Ulrich Steiner, Isabelle Guyon, ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition datasets and results, ChaLearn Looking at People workshop, ICCV, 2015.

Chalearn LAP – Apparent age estimation

- CVPR’16 - Face Analysis Workshop and Challenge
- ICCV’15 - Apparent Age Estimation and Cultural Event Recognition

- 8K images each displaying an individual,
- Labels consisting of,
  - Real and apparent age,
  - Gender,
  - Smile/no-smile,
  - Wearing accessories.

Input

Cropped Face

<table>
<thead>
<tr>
<th>GT Real</th>
<th>GT Apparent</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.00</td>
<td>28.84</td>
</tr>
<tr>
<td>30.00</td>
<td>34.30</td>
</tr>
<tr>
<td>25.00</td>
<td>30.11</td>
</tr>
<tr>
<td>31.00</td>
<td>33.05</td>
</tr>
<tr>
<td>29.00</td>
<td>34.84</td>
</tr>
<tr>
<td>18.00</td>
<td>26.16</td>
</tr>
</tbody>
</table>

• Eirikur Agustsson, Radu Timofte, Sergio Escalera, Xavier Baro, Isabelle Guyon, Rasmus Rothe, Apparent and real age estimation in still images with deep residual regressors on APPA-REAL database, FG, 2017. [FG 2017 one of the best papers award]
• Sergio Escalera, Junior Fabian, Pablo Pardo, Xavier Baro, Jordi Gonzalez, Hugo Escalante, Dusan Misevic, Ulrich Steiner, Isabelle Guyon, ChaLearn Looking at People 2015: Apparent Age and Cultural Event Recognition datasets and results, ChaLearn Looking at People workshop, ICCV, 2015.
Chalearn LAP – Faces of the world: accessories and gender recognition

- CVPR’16 - Face Analysis Workshop and Challenge
- 25,000 images collected from flickr

<table>
<thead>
<tr>
<th>Accessory</th>
<th>Train</th>
<th>Validation</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hat</td>
<td>1151</td>
<td>608</td>
<td>869</td>
</tr>
<tr>
<td>Headband</td>
<td>243</td>
<td>109</td>
<td>193</td>
</tr>
<tr>
<td>Glasses</td>
<td>1232</td>
<td>614</td>
<td>828</td>
</tr>
<tr>
<td>Earrings</td>
<td>770</td>
<td>389</td>
<td>592</td>
</tr>
<tr>
<td>Necklace</td>
<td>615</td>
<td>300</td>
<td>559</td>
</tr>
<tr>
<td>Tie</td>
<td>151</td>
<td>72</td>
<td>220</td>
</tr>
<tr>
<td>Scarf</td>
<td>256</td>
<td>137</td>
<td>256</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Train</th>
<th>Validation</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>2946</td>
<td>1691</td>
<td>4614</td>
</tr>
<tr>
<td>Female</td>
<td>3318</td>
<td>1361</td>
<td>3799</td>
</tr>
<tr>
<td>Smile</td>
<td>2234</td>
<td>1969</td>
<td>4411</td>
</tr>
<tr>
<td>No Smile</td>
<td>3937</td>
<td>1117</td>
<td>3849</td>
</tr>
</tbody>
</table>

Chalearn LAP – Apparent personality

• NIPS 2016 Demo

✓ CVPR’17 – Explainable computer vision and Job candidate screening competition
✓ IJCNN17, CVPR’17 – Job candidate screening coopetition
✓ ICPR’16 - Multimedia Challenges Beyond Visual Analysis
✓ ECCV’16 - Apparent Personality Analysis

• +10K 15s video,
• Big five personality traits,
• Video transcription.

Chalearn LAP – Sign gesture recognition

✓ ICCV 2017
Chalearn LAP – Real-fake emotion recognition

- ICCV 2017
- Recognition of fakeness of an emotion
- Recognition of fake/true specific emotion
- 6 basic emotions
- 100FPS

- **SASE-FE Dataset**

<table>
<thead>
<tr>
<th>Sets</th>
<th># Labels</th>
<th># Videos</th>
<th># Performers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>12</td>
<td>480</td>
<td>40</td>
</tr>
<tr>
<td>Validation</td>
<td>12</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>Test</td>
<td>12</td>
<td>60</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>12</td>
<td>600</td>
<td>50</td>
</tr>
</tbody>
</table>
The books of this innovative series collect papers written by successful competitions in machine learning. They also include analyses of the challenges, tutorial material, dataset descriptions, and pointers to data and software. Together with the websites of the challenge competitions, they offer a complete teaching toolkit and a valuable resource for engineers and scientists.

**Book: First NIPS ’17 Competition**

- Competitions volume chapter submission 15th January, 2018
- Book chapter acceptance notification 15th March, 2018
- Camera ready submission 15th April, 2018
- Tentative publication of the book 1st July, 2018

Info: [sergio.escalera.guerrero@gmail.com](mailto:sergio.escalera.guerrero@gmail.com)

ChaLearn Looking at People

http://chalearnlap.cvc.uab.es/
Thank you!

If you want to use codalab and need support, or you want to propose a challenge being sponsored by ChaLearn please contact us!

admin@chalearn.org

ChaLearn Looking at People
http://chalearnlap.cvc.uab.es/

Codalab
Open source: https://github.com/codalab/codalab-competitions
Public instance: https://competitions.codalab.org/