

## COMP-551: Applied Machine Learning

### Project 4: Reproducible Machine Learning

Due on December 4 or 6, in class (Sign up, Spotlight presentation)

Due on December 15, 11:59pm (Written report, code, public review)

**Background:** One of the challenges in machine learning research is to ensure that published results are reliable and reproducible. In support of this, the goal of this project is to investigate reproducibility of empirical results submitted to the International Conference on Learning Representations (iclr.cc). Papers submitted to the conferences are available for public review:  
<https://openreview.net/group?id=ICLR.cc/2018/Conference>

You should select a paper from this list, and aim to replicate the experiments described in the paper. The goal is to assess if the experiments are reproducible, and to determine if the conclusions of the paper are supported by your findings. You do not need to reproduce all experiments in your selected paper, for example the authors may experiment with a new method that requires more GPUs than you have access to, but also present results for a baseline method (e.g. simple logistic regression), in which case you could elect to reproduce only the baseline results. It is sometimes the case that baseline methods are not properly implemented, or hyper-parameter search is not done with the same degree of attention. You can implement algorithms from scratch, or use any existing toolbox or software, including code released by the authors, as long as you reference everything appropriately in your report.

The result of the reproducibility study should NOT be a simple Pass / Fail outcome. The goal should be to identify which parts of the contribution can be reproduced, and at what cost in terms of resources (computation, time, people, development effort, communication with the authors). Essentially, think of your role as an inspector verifying the validity of the experimental results and conclusions of the paper.

**Scope:** The project scope is open-ended. In terms of expected workload, it should be roughly twice the effort of project 1, and 50% more than project 2 or 3. If you are unsure about scope or topic, consult with the teaching team to validate your plan. Drop in during office hours, or make an appointment. Do this early! The final project should be completed in groups of three. You are strongly encouraged to work with different people for every project, but you are allowed to work with a previous team member if you prefer.

For more details on the challenge, including other participating institutions, cloud computing credits, etc.:  
<http://www.cs.mcgill.ca/~jpineau/ICLR2018-ReproducibilityChallenge.html>

**Requirements:** There are 5 mandatory components to submit (read instructions carefully!)

1. **Sign up** your team for the challenge:  
<https://docs.google.com/forms/d/1GAZnZWYW2suf6Z9polBITQvTvMJjkMy7CNyMapNkuY/edit?ts=59d53577>
2. **Spotlight presentation** (1 / team). Prepare a 3min. presentation (5-6 slides max) describing your project. This should clearly outline the target question, describe the methodology, and if possible, give preliminary results. Sign-up for a presentation slot:  
[https://docs.google.com/spreadsheets/d/1G\\_wGgR7leHvfr2TSri\\_IrMVZwXZGXgtx-nlik-4GSZo/edit#gid=0](https://docs.google.com/spreadsheets/d/1G_wGgR7leHvfr2TSri_IrMVZwXZGXgtx-nlik-4GSZo/edit#gid=0)  
*Note: There is 1 tab for Monday Dec.4 and 1 tab for Wednesday Dec.6., and 2 parallel sessions.*  
Submit slides here before your presentation time (you will need to sign up with a Google account):  
<https://drive.google.com/drive/folders/15AtV4cjE2Zlj5KgzG4vDm8QLkcN720Mp?usp=sharing>  
Evaluation criteria: Clarity and feasibility of reproducibility plan, level of detail, work to date.  
Weight: 10% of final project grade.
3. **Written report** (1 / team). The report should have at most 8 pages and be submitted on CMT. It should follow a standard conference format (see instructions from previous projects). The report should clearly present the target questions, a clear and well-motivated methodology, analysis and discussion of findings. Weight: 70% of final project grade.
4. **Code** (1 / team). Any code developed should be uploaded as supplementary material on CMT.

5. **Public review** (1 / team): Prepare and publish on OpenReview an executive summary (roughly 1 page) of your full report; be as detailed as possible on key findings. You are encouraged to include a link to your full written report. Make sure to support any statement with as much evidence as you can. Include a copy of the public review as an appendix to your written report (in the same pdf file) when submitting on CMT. The public review should also be submitted publically on (you are encouraged to get feedback from a TA or Herke or Joelle before posting):

<https://openreview.net/group?id=ICLR.cc/2018/Conference>

To submit your public review you will need to create an account and login on OpenReview, then find the paper you have studied and select "Public Comment".

Weight: 20% of final project grade.