Task Description

You can select a paper from the list of accepted papers from NeurIPS 2020, ICLR 2020, ICML 2020, ACL 2020, EMNLP 2020, CVPR 2020 and ECCV 2020, and aim to replicate the main claim described in the paper. The objective is to assess if the experiments are reproducible, and to determine if the conclusions of the paper are supported by your findings. Your results can be either positive (i.e. confirm reproducibility), or negative (i.e. explain what you were unable to reproduce, and potentially explain why).

Essentially, think of your role as an inspector verifying the validity of the experimental results and conclusions of the paper. In some instances, your role will also extend to helping the authors improve the quality of their work and paper.

Task scope

We recommend you focus on the central claim of the paper. For example, if a paper introduces a new RL learning algorithm that performs better in sparse-reward environments, verify that you can re-implement the algorithm, run it on the same benchmarks and get results that are close to those in the original paper (exact reproducibility is in most cases very difficult due to minor implementation details). You do not need to reproduce all experiments in your selected paper, but only those that you feel are sufficient for you to verify the validity of the central claim.

If available, the authors' code can and should be used; authors increasingly release their code and this is increasingly seen as an integral part of the publication process. Just re-running code is not a reproducibility study, and you need to approach any code with critical thinking and verify it does what is described in the paper and that these are sufficient to support the conclusions of the papers. Consider designing and running unit tests on the code to verify it works well and as described. Alternately, the methods presented can also be fully re-implemented according to the description in the paper. This is a higher bar for reproducibility that can take much more time, but may be helpful in detecting anomalies in the code, or shedding light on aspects of the implementation that affect results. In the end, what you choose to do will depend on your resources and how confident you want to be about the central claim of the paper.

Generally, a report should include any information future researchers or practitioners would find useful for reproducing or building upon the chosen paper. The results of any experiments should be included; a "negative result" which doesn't support the main claims of the original paper is still valuable.

We also strongly encourage you to get in touch with the original authors to seek clarification and make sure your reproducibility report fairly reflects on their research and work with them to improve it.

Proposed outcomes

- The goal of this challenge is **not** to criticize papers or the hard work of our fellow researchers. Science is not a [competitive sport](). Thus, the main objective of this challenge is to provide a fun learning exercise for newcomers in the Machine Learning field, while contributing to the research by strengthening the quality of the original paper.

- Participants should produce a Reproducibility report, describing the target questions, experimental methodology, implementation details, analysis and discussion of findings, conclusions on reproducibility of the paper. This report should be posted as a contributed review on OpenReview.
- 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).
- Participants should expect to engage in dialogue with original paper authors through the OpenReview site. Reproducibility Reports will be published at ReScience journal after peer review through OpenReview.

7