

Continuous Open Review

Introduction

At OpenAI, we are building a plan to improve the machine learning conference system. Conferences play two major roles that we feel are approximately orthogonal: they review and publish papers, and they are social gatherings where researchers meet to learn from each other. This document describes our plan for improvements to the reviewing and publication system. We have a separate plan for an Unconference that we hope will provide a new kind of

Request for feedback

We'd like to hear your thoughts about this proposed system. In particular, we'd like to know:

- Do you have a suggestion for the name for the new reviewing tool? (“Open Review” is taken, see openreview.net)
- Do you strongly disagree with any of our proposed actions or philosophies?
- Is there anything else you think we should plan to incorporate in this tool in order to:
 - Increase the representation of underrepresented groups of machine learning researchers?
 - Make the system more fair for junior, not-yet-established researchers?
 - Improve the quality of papers?
 - Reduce the amount of effort required for reviewers to yield good results?
 - Accelerate progress in the field?

Key Idea

We want to build a reviewing system that combines the ongoing back-and-forth process between authors and reviewers that occurs in a journal system with the speed of dissemination of research that happens on arxiv.

When progress in machine learning was slower, journals were appropriate. Now that progress is extremely fast, we feel that new reviewing procedures are necessary.

BASIC PROPOSAL

The basic idea is to build a web tool that associates arxiv papers with reviews.

One form this web tool might take is a browser extension. When visiting an arxiv page, the web extension

Readers can then browse arxiv through our wrapper, rather than through arxiv directly. This way, endorsements and criticisms of papers appear automatically as they browse.

All reviews that are ever prepared will become permanently public and permanently associated with the paper.

Anyone can request that a paper be reviewed.

Reviewers may choose to be anonymous or not. Reviews are considered mini-publications that may be cited.

Reviews will be linkable and shareable.

In addition to writing reviews that comment on the quality of the paper, it will be possible to write brief summaries of papers, similar to what Hugo Larochelle does with Evernote.

Reviews can contain specific tags like:

- Contains new algorithm
- Contains new theorem
- Achieves new state of the art
- Fails to cite important related work
- Experimental evaluations are appropriate
- Experimental evaluations are not appropriate (weak baseline, dataset that not many people publish on, etc.)

POSSIBLE EXTENSIONS

These features take more implementation effort, but are probably desirable:

- Authors are given the option to release a paper anonymously (with their identity known to the server but not to the public). It appears on arxiv without any names or affiliations. The authors can choose to have the server reveal their names after the paper has accumulated some reviews. This feature would require more work because arxiv does not currently support anonymous submissions, so it is not just an annotation applied on top of arxiv.
- Reviewer economy: each reviewer account will be associated with a balance of reviewing points. To request a review of a paper, it is necessary to spend one reviewer point. To earn a reviewer point, one must review a paper. (Or similar system. Obviously this requires a mechanism for initializing the system, like having each new user begin with a loan of n reviewer points from a fictitious bank)
- In general, we think it will be tricky to tune the review quality system—this is something that companies like reddit and Quora already need to do in a slightly different context and it is generally regarded as challenging. We don't expect to get the system exactly right in this initial proposal, but we plan to adapt the system to challenges as they arise in practice. Some measures we think might be beneficial for the future include:
 - Duplicate detection: when using the web tool to view an arxiv paper, the wrapper will automatically identify other arxiv papers that are near-duplicates and show their reviews too. This is to prevent authors from trying to escape bad reviews by resubmitting the paper.
 - Reviewer reputation system: while most reviews will probably be anonymous, reviewers should have a public profile that displays their general review reputation. Reviewers (not authors) will be asked to do mini-reviews of reviews, by just doing a simple upvote / downvote of reviews. This earns you stackoverflow style karma. Karma must be implemented with care to preserve reviewer anonymity, and will probably require features like delayed and partially randomized updates to the publicly visible karma page.

- To prevent people from trading reviews, we might need a feature like preventing reviewer cycles. In turn, this might require a feature that randomly forbids some reviewer interactions in order to preserve anonymity.
- Mechanisms for verifying the real-world identity of authors or reviewers.
- Mechanisms for banning accounts engaged in malicious behavior.

GOALS AND PHILOSOPHY

We believe this will improve over the existing reviewing system for the following reasons:

- Compared to journals: papers appear immediately.
- Compared to arxiv without annotations: papers are reviewed, so readers get an idea of whether to trust them or not.
- Compared to conferences:
 - Authors do not need to rush to prepare papers by a hard go / no-go deadline so papers should have higher quality of presentation.
 - Reviewers can be added to papers gradually over time and authors can respond to reviewer criticisms gradually over time, like in the journal reviewing system.
 - Because there is no hard go / no-go paper acceptance deadline, interested and relevant reviewers can be recruited gradually over time. This should improve review quality, because in conferences, reviewers are often assigned to review papers that are outside of their background, due to the pressure to do all of the community's reviewing simultaneously in a short reviewing cycle.
- General advantages:
 - Currently, if a paper containing a major error is posted on the web, then rejected from a conference or journal, there is no warning label associated with the paper. Subsequent authors usually cite the paper out of a sense of obligation to cite all relevant work, despite the major error. With the rising popularity of machine learning yielding more production of low-quality machine learning papers, and the advent of arxiv as a major web publishing venue, this has become a significant issue.
 - A small number of malicious reviewers can currently sink a paper. If an author submits a paper to a conference, and gets a reviewer who wants to be known as the sole expert in the field of that paper, the malicious expert can vote the paper down and then that's the end of it. With continuous review, the author could continue to expend review points to acquire more reviews. If the paper is actually good, eventually the supply of malicious reviewers would be exhausted and the paper would start to accumulate positive reviews.
 - Existing attempts at open reviewing often fail to be transparent in significant ways. For example, reviews and scores may be made public but the real decisions may be made by area chairs, whose decision making process is not documented or released. Reviews may be released only for accepted papers. Reviews of rejected papers are more important, because they show the flaws in bad work on arxiv, and because that is where malicious reviewers sink good work for unfair reasons. In some cases, reviews are published after

the review process has completed, but reviewers are given the option of editing their review prior to publication.

- The workload on reviewers will be reduced, because people will only be expected to do as much reviewing as they require in the reviewing economy.
- The demand for reviewing will be reduced, because fewer people will request reviews than currently submit to conferences. The current conference review system gives authors an incentive to submit bad work to multiple conferences hoping that noise in the system will eventually result in acceptance (a “statistical fishing expedition”). The new system will permanently associate all reviews with a paper, so authors will have an incentive to wait to request review until the work is in a good work.