How to ML Review — A brief Guide

How to ML is a series of non-technical ML research advice posts by @j_foerst and @rockt. If you like this, check out "How to ML Rebuttal — A Brief Guide", "How to ML Paper – A Brief Guide", follow @HowTo_ML for future posts, and share on X.

First off: The review process is at the core of the community so first off, thank you for taking it seriously and for taking time! Reviewing is an extremely important part of the academic community. Area chairs will take note of reviewers who do a fantastic job (see for example Outstanding Reviewer awards by conferences like ICLR) but also remember reviewers who submit low-quality reviews or vastly overshoot deadlines without communicating delays. Secondly: These are general guidelines based on our experience reviewing, meta-reviewing, senior area-chairing etc. Wherever they conflict with the guidance provided by the conference or journal, the latter official sources take precedence.

On a high level, we believe the best approach to reviewing is to temporarily see yourself as a collaborator to the authors' work who is particularly critical and who is looking out for holes in the authors' arguments or empirical work to help them improve their paper. Consequently, as you become a better reviewer, you will also get better at criticising your own work and that of your colleagues, and thus do better research. That is why we encourage our PhD students to learn to take part in reviewing early on.

Step 0 (optional but highly suggested):
- Read the How to ML Paper Guide – while not every author follows this, having a framing for how to write and therefore also parse ML papers is quite useful.

Step 1 (Accepting to Review):
- You just received the email inviting you to review. First off, *congratulations*! This means someone values your judgement and perspective highly enough to ask for your input in this crucial process.
- Before you click "accept":
  - Take a look at the reviewer guidelines and *timelines.*
  - Open your calendar and see how the reviewing period and discussion period fit into your schedule. Reviewing takes time.
  - As a rule of thumb, we recommend budgeting around 6-12 hours for each paper that you accept to review, with more time needed if this is your first time reviewing. Note, this number will naturally differ by field – e.g. theory papers usually take more time.
- Once you decide to commit, enter *blockers* in your calendar to allocate time for reviewing with plenty of margin of error ahead of the deadline.
- Make sure your conflict of interests are up-to-date in the reviewing tool and that you entered the correct "maximum number of papers" if possible.

Step 2 (bidding):

- This part is often overlooked, but crucial. Every minute you spend here can potentially avoid a lot of time sinks down the line. Many conferences (like ICLR, ICML and NeurIPS using OpenReview) also automatically match papers based on your prior publications. Make sure to keep the list of your own co-authored papers where you have relevant expertise up-to-date!
- When you bid on papers, there are a few key questions:
  - Are you an expert in the area of the paper? Specifically, are you familiar with the state-of-the-art in the field, relevant related work, common approaches, etc?
  - Are you familiar with the background knowledge required for the specific approach taken by the authors? For example, if the paper is using Variational Autoencoders for multi-agent learning, it is very helpful for you to know about VAEs.
  - Does that paper seem exciting to you? It's a lot easier to spend a large amount of time on reviewing a paper if we genuinely care.

Step 3 (Reading The Papers):
- Allocate about 1 hour for the first full parse of each paper that you are reviewing.
- In general, we recommend reading (and thus also reviewing) papers in multiple phases (following the excellent guide by Adler, M. J., & Doren, C. V. "How to Read a Book"). First, a *skim read* where you try to get the gist of the paper by reading the title, abstract, and briefly look into the methods and main results of the paper. Second, a *full analytical reading* of the paper where you try to understand every single sentence and claim made by the authors. This is also where you would start to form your opinion about the paper — are the claims following logically? Does the empirical evidence support the claims? Lastly, some papers require even more work to review as it might not be possible to assess the novelty, impact or accuracy of the papers claims based on the submitted paper alone. You might have to do some syntopical reading, comparing the paper under review with other related papers in the field. Often, this will be easy for you since you are already an expert in the field and know the most relevant prior work.
- Type out notes during the skim read and the analytical reading of the paper. Close all other tabs in your browser and turn off your phone. Time to focus. Even better: print out the paper!
- *Example note-keeping:*
  - *Key contributions:* What does this paper claim to do?
  - *Key Novelty*: What is the closest alternative from the prior art and how does this differ?
  - *Key Evidence 1*: What evidence do the authors provide that their method performs well?
    - If empirical:
      - Is the method tested on a sufficiently broad set of benchmarks?
      - Are the results statistically significant and does the evaluation protocol follow best practices (reporting confidence intervals, significance testing)?

- - Do the baseline numbers match work published in prior state-of-the-art methods in this field?
    - Are the compute / data requirements stated and comparable across methods and benchmarks? If not, is this discussed in the paper?
  - If theoretical:
    - Are the proofs correct?
    - Are they easy to follow?
- *Key Evidence*:
  - What evidence do the authors provide that their method performs as intended?
    - Part 1 of the evidence showed the method beats SoTA.
    - In contrast, this is about *Ablations* and analysis.
    - The expectation here is that authors show *ablations* which verify that their key contributions make the difference, keeping all else the same.
    - Note: A common issue here are hyperparameters. For example, if the authors tuned their parameters for the method, you need to make sure that they re-tuned them fairly for the ablation (i.e. the same as the method, but one component removed).
- *Key Questions / Issue*:
  - Where do you see possible problems with the line of argument?
  - Which parts of the paper are unclear?
  - Is there any notation that is inconsistent or not introduced?
  - Where do you see possible improvement?
  - Given the information provided in the paper, would you be able to replicate it?
- It is useful to read all of the papers in your batch before proceeding to the next step. This will help you calibrate expectations, especially if you are new to reviewing.

Step 4 (Writing the Review):
- Take another look at the reviewer guidelines and the actual review form.
- Personally, I recommend copying over the questions from the review form into a google doc and using this as your editor. It makes editing, spell checking etc easier and also avoids losing work if the site crashes or times out.
- There are a couple of very high quality reviews here: https://openreview.net/forum?id=B1xm3RVtwB – please take a look at them for inspiration. Note, the original reviewers are fairly negative, but very objective. The paper went from something like 3/3/3/ to 8/8/8 based on the authors addressing the concerns and the reviewers engaging.
- Hopefully most of the questions in the review form are easy to answer now based on the notes you took while reading the paper.
- Things to focus on in your writing and judging:

- ○ Be *constructive*: The goal of the review process is both to identify good papers and to help authors improve their work, so constructive pointers are super useful.
- ○ Be *specific*: It is *not* ok to say "many other papers have done this before", without providing references to the specific papers. It is *not* ok to say "the writing is unclear", without providing specific instances and suggestions for improvement.
- ○ Be *reasonable*: You might disagree with the premise of an entire established subfield or problem setting, but rejecting individual conference papers is not the right venue for addressing this.
- ○ *Justify your score / recommendation*: Say what recommendation you are going to give and why.
- ○ *Articulate expectations*: If there are specific issues that you would like to see addressed or additional experiments that are missing before you can improve your score, please state them clearly and say that you would then be happy to improve your score.

Step 5 (Scoring):
- ● Take another close look at the reviewer guidelines – what constitutes a borderline paper? What is a clear rejection? What is a clear acceptance?
- ● Be decisive: If possible, after having made a substantial attempt at following the authors' arguments and presentation of empirical evidence (if any), form a justified opinion about the merits and weaknesses of the work, and whether the paper should be accepted. If you can, avoid borderline ratings.
- ● Often this comes down to balancing different qualities of a paper, some of which are more subjective than others.
  - ○ Does the paper make meaningful progress
- ● Common Reasons for rejection:
  - ○ Fundamentally technically wrong
  - ○ Misleading
  - ○ Related work was missed which already did what the paper claims to be doing as a novel, key contribution. You need to be very specific in your claims and double check that the key contribution is the same.
- ● Not valid reasons for rejection:
  - ○ Poor English and grammar issues (unless they make it impossible to judge the merits of the paper)
  - ○ Related work was missed
  - ○ Anything else that can be addressed with writing changes (though it's recommended to ask for these changes in the review)
- ● Common Reasons for acceptance:
  - ○ New method with strong empirical results in an established domain (i.e. prior conference papers published)
  - ○ Novel theoretical results
  - ○ New analysis method and insights in a relevant area of Machine Learning
- ● Clearly, this leaves a lot of room for the middle-ground where judgement comes in. Here it is important to trade-off the *realised potential* of a paper (strong results) versus the

potential for setting new directions and creating downstream impact. Check you were specific in your review writing when weighing off the two to justify your scoring.
● Lastly, keep an open-mind about changing your score based on the author's rebuttal and discussion period (see below).

Step 6 (Discussion Period – if there is one):
● First-off, *please engage.* Authors put a huge amount of work into rebuttals and into addressing reviewer concerns, so ignoring the authors' response during the discussion phase is quite bad.
● Not all engagement is equal, though. For example, when there is acknowledgement of the author rebuttal required by the process, don't simply say "I saw the response and will keep my score". Instead, be specific about which parts of your concerns were addressed and which ones were not, including why.
● Even increasing your score is often not sufficient. For example, if you say that X,Y,Z, would need to be addressed to raise your score and the authors address X,Y,Z, then increasing the score from a "borderline reject" to a "borderline accept" might not be enough. If all the concerns were addressed, why not a better score? Make sure to either increase further or add rationale why the only slightly updated score is now accurate.
● Engage in the discussion with other reviewers and area-chairs. Do not wait for the area-chair to initiate a discussion. Ideally, be proactive. Ask other reviewers about clarifications. Call out biases. If you strongly believe a paper should be accepted or rejected, champion the respective position, but make sure you can back up your position with clear arguments.

Frequently Asked Questions:

What happens if you receive a paper that you have reviewed in a prior conference? What if the authors submitted the exact version without any improvements?
➔ flag to AC and call out authors for not having addressed the concerns you already voiced in a previous round of reviewing.

What should I do if I have an emergency and can't submit my reviews in time?
➔ tell AC well in advance, and give them a clear indication how much more time you will need or whether they should find an emergency reviewer to replace you.

Can I share the ideas of a cool paper that I am reviewing with my friends?
➔ no, generally papers you are reviewing should be kept strictly confidential.

Should I Google the paper I am reviewing to see if it is already published?
➔ no, you should not take actions that would break the anonymity of the authors. However, if you accidentally happen to find out the names of the authors (e.g. because they post the paper on X), that is usually fine (see below).

What happens if I accidentally find out who the authors are for a paper that I am reviewing?

➔ This is usually fine, but it's usually still a good idea to inform the area chair with a private message. You absolutely have to inform the AC if it turns out that there was a conflict of interest between you and the authors, which should have been avoided in the paper matching phase in the first place.