How to ML Paper - A brief Guide

Feel free to comment / share and good luck with NeurIPS!

Canonical ML Paper Structure

Abstract (TL;DR of paper):

X: What are we trying to do and why is it relevant?
Y: Why is this hard?
Z: How do we solve it (i.e. our contribution!)
1: How do we verify that we solved it:
1a) Experiments
1b) Theory

Introduction (Longer version of the **Abstract,** i.e. of the entire paper):

X: What are we trying to do and why is it relevant?
Y: Why is this hard?
Z: How do we solve it (i.e. our **contribution**!)
1: How do we verify that we solved it:
1a) Experimental results
1b) Theory
Extra space? Future work!
Extra points for having *Figure 1* on the first page

Related Work:

Academic siblings of our work, i.e. alternative *attempts* in literature at trying to solve the same problem.
Goal is to "*Compare and contrast*" - how does their approach differ in either *assumptions* or method? If their method is *applicable* to our **problem setting** I expect a comparison in the experimental section. If not there needs to be a clear statement why a given method is not applicable.
Note: Just describing what another paper is doing is *not enough.* We need to *compare and contrast*.

Background:

Academic Ancestors of our work, i.e. all concepts and prior work that are required for understanding our method.
Includes a subsection **Problem Setting** which formally introduces the problem setting and notation (**Formalism**) for our method. Highlights any specific assumptions that are made that are unusual.

Method:

What we do. Why we do it. All described using the general **Formalism** introduced in the **Problem Setting** and building on top of the concepts / foundations introduced in **Background**.

Experimental Setup:

How do we test that our stuff works? Introduces a specific *instantiation* of the **Problem Setting** and *specific implementation details* of our **Method** for this **Problem Setting.**

Results and Discussion:

Shows the results of running **Method** on our problem described in **Experimental Setup**. Compares to baselines mentioned in **Related Work.** Includes statistics and confidence intervals. Includes statements on hyperparameters and other potential issues of fairness. Includes *ablation studies* to show that specific parts of the method are relevant. Discusses limitations of the method.

Conclusion:

We did it. This paper rocks and you are lucky to have read it (i.e. brief recap of the entire paper). Also, we'll do all these other *amazing things* in the future.

Other Advice

Start with an *outline* rather than full text. Each *line* will correspond to *one paragraph* in the final version. It is much easier to change the outline of a building *before building it.* This is a great point in time to have conversations with others if you are unsure.

Next, expand the outline, but keep the summary text as Latex comments ahead of every paragraph. This will a) keep you on track and b) make it easy for anyone providing feedback to quickly see what the overall flow is.

Extremely common writing pitfalls and other advice (print this out and tick off?):

- ☐ Passive voice - unclear why, but this is a very common mistake. Passive gets way overused (e.g. here?). It's clunky and obfuscates who did what. Avoid it if it can be avoided.
- ☐ Be extremely clear on ***contributions***. Never blur the lines between what had been done before and what you did.
- ☐ Be consistent with tense. Avoid switching at all costs, also avoid using future tense if it can be avoided:
  - ☐ Eg. "In Section 3 we ~~will~~ show"
- ☐ Avoid filler words at *all cost.* Think about what you are trying to say and then say it, nothing else. Common filler words are " can ", " In order to ", and many others.
  Example:
  "The Bank Loan problem can be reformulated as a special subset of the contextual bandit problem" =>
  "The Bank Loan problem is a special instance of a contextual bandit problem"
- ☐ Once you have written the initial text, try to *delete around ⅓ of the words.* That's typically how much "fluff" there is.
- ☐ Please use correct quotation marks in Latex ``correct quotation'' (copy-paste this if unclear).
- ☐ Use *"*\citet*"* when authors are part of the sentence, e.g. *"*\citet{foerster2016learnig} show *..",* and *"*~\citep*"* otherwise, e.g *"..* recent work~\citep{foerster2016learnig}*"*.

- [ ] Cite any claim that is not supported by your experiments and avoid grandiose language or overly broad claims - it usually makes it easy to attack the paper for no good reason
- [ ] Don't leave writing the paper until the last minute. Aim for a *complete draft a week before the deadline.*
- [ ] Enable *change-tracking* in Overleaf and share directly with the email addresses of your collaborators. That way it's in their UI.
- [ ] Introduce any acronym *before* using it.
- [ ] Avoid *synonyms* at all costs.
- [ ] Only introduce *symbols* and *acronyms* that you use in the paper.
- [ ] What is **bold** and what is *italic*? Up to you, but be consistent.
- [ ] Avoid ~~R~~random ~~C~~capitalisation (RC), even for method names and when introducing acronyms.
- [ ] Avoid anthropomorphisms ("knowledge" etc) of AI algorithms.
- [ ] Avoid subjective claims - usually *adjectives* are red flags.
- [ ] "On the other hand" can't come without "On the one hand".
- [ ] Watch out for repetition of words within a single paragraph.
- [ ] Use *simple language* when possible. Avoid rare words or sounding "fancy''. For plenty of scientists (like myself) English is not the first language, don't make life hard for them.
- [ ] Footnotes should be after "." and "," (credits to Oana).
- [ ] Never *copy-paste* from other papers, unless you are verbatim quoting something. It's much easier (and more ethical!) to write something from scratch than to try to modify something until it looks different.

Last not least - communicate plenty with all authors (i.e. at least daily for the last week) to stay on track and *have fun*!!

PS: Nothing here is binding but I think it makes it much easier for everyone if we stick to a basic structure when writing papers. Think of it like a broad convention that allows readers to quickly process papers.

PPS: These broad best-practice suggestions are the result of writing papers with a fantastic set of mentors, students and other collaborators - all credits go to them!

Comments / questions? Email jakob foerster at eng dot ox dot ac dot uk or comment on Twitter.