

Branch: develop ▾ cs224u / projects\_paper.md

[Find file](#) [Copy path](#) andreionutdamian update

d3087d3 3 minutes ago

[1 contributor](#)

228 lines (102 sloc) 21.3 KB

[Raw](#) [Blame](#) [History](#)   

## Final paper

The final paper is where it all comes together. We assume you're shooting for a great grade. We encourage you to think beyond that as well. Some motivations for thinking big:

- Many important and influential ideas, insights, and algorithms began as class projects.
- Getting the best research-oriented jobs will likely involve giving a job talk. Your project can be the basis for one.
- You can help out the scientific community by supplying data, code, and results (including things that didn't work!).

It's worth reiterating a central point from [the evaluation methods notebook](#): We will never evaluate a project based on how "good" the results are. Publication venues do this, because they have additional constraints on space that lead them to favor positive evidence for new developments over negative results. In this course, we are not subject to this constraint, so we can do the right and good thing of valuing positive results, negative results, and everything in between. Thus, we will evaluate your project on:

- The appropriateness of the metrics.
- The strength of the methods.
- The extent to which the paper is open and clear-sighted about the limits of its findings.

## Formatting

The paper should be 8 pages long, in ACL submission format and adhering to ACL guidelines concerning references, layout, supplementary materials, and so forth. [Here are the LaTeX and Word templates for the current ACL style.](#)

## Suggested paper organization

Papers in the field tend to use a common structure. You are not required to follow this structure, but doing so is likely to help you write a paper that is easily understood by practitioners, so it's strongly encouraged.

### Abstract

Ideally a half-column in length. It's good to give some context for the work right at the start, define the current proposal and situate it in that context, summarize the core findings, and close by identifying the broader significance of the work. The "General reasoning" section of your experiment protocol is likely to provide good material for the abstract.

### Introduction

1–2 columns. This is an extremely important part of the paper. In this section, the reader is likely to form their expectations for the work and begin to form their opinions of it. The introduction should basically tell the full story of the paper, in a way that is accessible to most people in the field:

1. Where are we? That is, what area of the field are you working in? Answering this question is important for orienting the reader.
- 2.

What hypothesis is being pursued? It's a good sign if you have a sentence that starts with a phrase like "The central hypothesis of this paper is ...". You don't need to be this explicit, but, on the other hand, this is a way of ensuring that you don't end up saying only vague things about what your hypothesis is. Also, being direct about this can expose a lack of clarity in your own thinking that you can then work through.

3. What concepts does your hypothesis depend on? You can't require your reader to fill in the gaps. Try to place all the building blocks of your hypothesis in a way that supports the hypothesis itself. Sometimes this material is best given after the hypothesis statement, but very often it needs to be given before, so that the hypothesis itself makes sense.
4. Why *this* hypothesis? What broader issues does it address? This will provide further context for your ideas and help motivate your work.
5. What steps does the paper take to address your hypothesis?
6. What are the central findings of the paper, and how do they inform the core hypothesis?

### Related work

1.5–2 columns. This section can draw heavily on your lit review. A good strategy for this section is to first organize the papers you want to cover into general groups that relate to your own work in important ways. For each such group, articulate its thematic unity, briefly discuss what each paper achieves, and then, crucially, relate this work to your own project, as a way of providing context for your work and differentiating it from prior work. In this way, you carve out a place for your own contribution.

### Data

Length highly variable. This section should describe the properties of your dataset. This typically means giving some actual examples (or descriptions of them, if the examples are very long) as well as quantitative summaries (e.g., number of examples, number of examples per class, vocabulary size, etc.). With this discussion, you are trying to motivate your dataset choice, convey what your task is like, and build up context for your later analyses. If you are describing a new dataset that you created, then you should plan to devote a lot of space to describing and motivating the data collection methods you used. [The Datasets section of your experiment protocol](#) is likely to provide good material.

### Your models

Length highly variable. Your experiment protocol should provide basic descriptions that you can expand and polish for this section.

If you are proposing an original model or exploring a complex, non-standard model from another paper, then this section will have to be quite long and detailed. With luck, your preceding "Related work" and "Data" sections provide some conceptual support for what you are doing.

If you are just comparing familiar models, then this section might be shorter, but it is still crucial.

To the extent possible, it's good to separate your model description from particular choices that you made for your experiments, as those are really part of the experimental design.

### Experiments

Length highly variable. This section should explain how the data and models work together for your experiments. The reader should get a clear picture of which models were evaluated, how they were trained, how the data were pre-processed and subdivided for the experiments, which metrics were used, and what the experimental outcomes were. If there are a lot of details to cover, it's probably best to move them to appendices, so that this section conveys just what is needed in order to follow the argument. The appendix can then, correspondingly, be a deep dive into all the nitty-gritty details.

### Analysis

Length highly variable. This section is more open-ended. First and foremost, it should say what the experimental results mean. It is often fruitful to support this core conclusion with error analyses and qualitative trends in the predictions that provide deeper insights into where your favored model is succeeding and failing.

## Conclusion

A half-column in length. Since NLP papers are short, the conclusion need not be long or detailed. It's probably best to briefly summarize what the paper did and why, and then try to articulate the broader significance of the work, perhaps looking ahead to expanding its scope. Some papers include a lot of analysis in the conclusion, which is fine, but I would advise that you put that material in a separate section, so that the conclusion works more as a succinct encapsulation of the paper itself, akin to the abstract but perhaps with more technical content, since it can build on the paper itself.

## Required authorship statement

At the end of your paper (after the 'Acknowledgments' section in the template), please include a brief authorship statement, explaining how the individual authors contributed to the project. You are free to include whatever information you deem important to convey. For guidance, see the second page, right column, of [this guidance for PNAS authors](#). We are requiring this largely because we think it is a good policy in general. This statement is required even for singly-authored papers, because we want to know whether your project is a collaboration with people outside of the class. Only in extreme cases, and after discussion with the team, would we consider giving separate grades to team members based on this statement.

## Advice on scientific writing

Here are some really nice pieces of writing about scientific writing:

- [Novelist Cormac McCarthy's tips on how to write a great science paper](#)
- [David Goss: Some hints on mathematical style](#)

## Some exceptionally well-written NLU papers

These papers are stand-outs for me not just because of the proposals they make, but also because the writing is great.

- Keith, Katherine and Brendan O'Connor. 2018. [Uncertainty-aware generative models for inferring document class prevalence](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 4575–4585. Brussels, Belgium: Association for Computational Linguistics.
- Pennington, Jeffrey; Richard Socher; Christopher D. Manning. 2014. [GloVe: global vectors for word representation](#). In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. Doha, Qatar: Association for Computational Linguistics.
- Peters, Matthew E.; Mark Neumann; Mohit Iyyer; Matt Gardner; Christopher Clark; Kenton Lee, and Luke Zettlemoyer. 2018. [Deep contextualized word representations](#). In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, 2227–2237. Association for Computational Linguistics.
- Zettlemoyer, Luke S. and Michael Collins. 2005. [Learning to map sentences to logical form: structured classification with probabilistic categorial grammars](#). In *Proceedings of the Twenty First Conference on Uncertainty in Artificial Intelligence*.

And check out this amazing contribution: [The Annotated Transformer](#) (Alexander Rush).

## Some CS224u papers that become publications

Here is a selection of recent CS224u papers that evolved into published work:

- Badlani, Rohan; Nishit Asnani; and Manan Rai. 2019. An ensemble of humour, sarcasm, and hate speech for sentiment classification in online reviews. *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT)*. Hong Kong: Association for Computational Linguistics.
- Benavidez, Susana and Andy Lapastora. 2019. Improving hate speech classification on Twitter. *Proceedings of LatinX in AI Research*. Vancouver.
-

Kolchinski, Y. Alex and Christopher Potts. 2018. [Representing social media users for sarcasm detection](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 1115-1121. Brussels, Belgium: Association for Computational Linguistics.

- Jiang, Hang; Yuxing Chen; Haoshen Hong; and Vivek Kulkarni. 2020. DialectGram: Automatic detection of dialectal variation at multiple geographic resolutions. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2020*. New Orleans: Association for Computational Linguistics.
- Li, Lucy and Julia Mendelsohn. 2019. [Using sentiment induction to understand variation in gendered online communities](#). In *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*, 156–166. New York: Association for Computational Linguistics.

I'm being careful not to call these "published CS224u papers" because all of them went through substantial revisions before they were accepted. Some got shorter while retaining their original scope, whereas others grew in scope and got much longer.

I should emphasize also that these papers all exceeded our expectations for course projects even when they were submitted, and they matured afterwards, so the above are not examples of what we're expecting by the end of the course itself. These are more like things to aspire to longer term.

## Beyond the final paper

You might be thinking about continuing to work on your final project after the class ends with the goal of publishing it. The goal of this section is to give you a sense for what that process is like, and to offer some tips on navigating it.

### Conference submissions

My focus is on conference submissions because NLP (like most areas of AI) is a conference-driven field in the sense that papers published at the top conferences are akin to journal papers in other fields. In NLP, the major exception to this is the journal *Transactions of the ACL*, which functions like a regular journal but (in a nod to the importance of conferences) also gives published authors the opportunity to present the work at one of the major ACL conferences.

My own view is that the top conferences in NLP are ACL, NAACL, and EMNLP (and that TACL is on par with these or soon will be). There might once have been real differences in prestige between these conferences, but those differences have disappeared. People submit their work as soon as possible after they feel it is ready, so the choice of these conferences comes down to the time of year.

COLING, CoNLL, EACL are also excellent conferences in NLP that might be seen as one step down from the above. Work from these conferences is aggregated in the ACL Anthology, which further reduces the importance of initial venue as times goes on.

Most NLP conferences host a series of workshops focused on specific areas of NLP and adjacent fields, and many of these workshops have proceedings volumes that are included in the ACL Anthology as well. Workshop papers are generally less prestigious than conference ones, but workshops are often more rewarding presentation venues than the main conference. The audience is full of people who are invested in the topic, the setting is more intimate, and a sense of community develops over the course of the workshop.

### The submission process

Here's what the submission process typically looks like in NLP:

1. The ACL conferences have adopted a uniform policy when it comes to posting work on the Web (including preprint servers like arXiv). The full policy is [here](#), and you should review it before submitting work for peer review to make sure you're abiding by it. The heart of the policy is that there is a window, starting one month before the conference deadline and extending through to the notification period, in which submitted work cannot be posted. The intention is to avoid influencing reviewers and compromising the double-blind reviewing process.
2. Assuming you're okay with regard to the submission policies, you submit your paper. Most conferences have a long paper option (8 pages) and a short paper option (4 pages). Sadly, the dynamic is that this means there can't really be, for example, 6 pages papers – they would look too slim for a long paper but violate the rules for a short paper.

3. At submission time, you separately choose area keywords and perhaps add some other similar metadata. This information will be used by the program committee when figuring out how to assign papers to specific reviewers, so it is worth thinking seriously about the choices here. If you can, try to find a veteran NLP author to get their advice on strategy.

Before submitting, make sure that your paper conforms to the conference's formatting requirements, so that it isn't desk rejected (i.e., rejected without peer review).

### The reviewing process

1. Once all the submissions are in, reviewers scan over long lists of titles and abstracts, and bid on which they want to review, which they will not review, which they would review if they had to, and which they can't review because of a conflict of interest. Realistically, the title is probably the primary factor in bidding decisions, since reading all the abstracts at this stage would be a huge task. Thus, when you craft your title, you should be thinking of these harried reviewers as a primary audience!
2. Once the bids are in, the program chairs assign reviewers their papers, presumably based in large part on their bids.
3. Reviewers read the papers, write comments, and supply ratings. The forms change from year to year, sometimes radically, but [the 2019 ACL form](#) looks pretty representative of what I tend to see as a reviewer these days. It is extensive, reflecting the seriousness of conference reviewing in the field.
4. Once the reviews are in, there is often a brief window in which the authors can respond to the reviews, with the goal of correcting mistakes and perhaps influencing the reviewers to adjust their scores. At this stage, the program chair might stimulate discussion among the reviewers about disagreements, the claims in the author response, etc. At this stage, all the reviewers see each other's names, which helps contextualize their reviews and creates some accountability.
5. Finally, the program committee does some magic to arrive at the final program based on all of this input. The procedures they use vary a lot and take a lot of factors into account.

### On giving good talks

Suppose, via a mixture of skill, hard work, and luck, your paper was accepted for publication. Very often, this comes with the requirement that you give a talk at the conference. You might also give talks in academic and industry settings – they are an important avenue for the dissemination of ideas.

#### Practical tips for giving talks

Here are some mundane things that you should attend to before taking the stage:

- Turn off any notifications that might appear on your screen.
- Make sure your computer is out of power-saver mode so that the screen doesn't shut off while you're talking.
- Shut down running applications that might get in your way.
- Make sure your desktop is clear of files and notes that you wouldn't want the world to see.
- If using PowerPoint / Keynote / Google Slides, have a PDF back-up just in case.
- Projectors can fail; always be prepared to give the talk without slides.

#### Talk contents

In general, the talk can mirror the structure of the paper. This is at least a good starting point, and then you can depart from it if you see a chance to make things clearer or add presentational flourishes.

However, talks have to be much less technical than papers can be. The emphasis should be on contextualizing the ideas and conveying the high-level narrative of what the paper does. You might begin from the premise that you'll have no equations or formulae, and then add them only where they are crucial and where you are certain you can devote sufficient time to presenting them.

Fundamentally, you're more likely to connect with the audience (and persuade them to read your paper) if you tell a clear and interesting story. There is a show-biz element to this!

The logician Patrick Blackburn (who gives wonderful talks) gave a (wonderful) talk years ago on how to give good talks. His central question was, "Where do good talks come from", and he answered with: "Honesty. A good talk should never stray from simple, honest communication." I think of this all the time, not just when preparing talks, but also when writing, creating teaching materials, and presenting ideas in meetings.

### The importance of practice

Practice is key. The more you practice, the more you'll hone your narrative and ensure that you get through it in the right amount of time. If you can recruit friends to listen, that's always beneficial. Friends who don't work in NLP can give especially valuable advice because they are less likely to unknowingly fill in gaps in your presentation.

It's also very helpful to record yourself giving the talk, because you'll learn a lot about what works and what doesn't by hearing yourself go through the material. It can be painful to listen to yourself, but it's worth it!

### The discussion period

Many conferences have a discussion period following each talk. These are sometimes called "question periods", but people are just as likely to make statements and offer criticisms as they are to ask questions, so "discussion period" is more apt! The discussion period is probably the most intimidating part of giving a talk. Giving the talk itself is nerve-wracking, but at least you're in control. In the discussion period, anything can happen!

The discussion period *should* be a chance for the audience to gain a deeper understanding of your ideas. When the entire question period has this aim, it is a joy. But sometimes other things happen. You'll encounter hostile questioners, self-serving questioners, and confused but adamant questioners. Whomever you encounter, and whatever they say, try to pause for one second before answering each question, even if you know exactly how to respond. It helps to set the right kind of conversational pace, and it signals that you're listening carefully.

Most questions won't make total sense to you. Your questioner doesn't know the work all that well, so try to be sympathetic. In general, you'll be a hit if you can warp every question you get into one that makes sense and leaves everyone with the impression that the questioner raised an important issue.

Try never to say, "I don't know" and leave it at that. When floored, consider instead saying, "I have no idea, but what about ..." and pushing the discussion in a new direction. This will make you look better, and it will enrich the discussion period itself.