CS224n: Project Proposal Instructions

Last updated on February 3, 2023

Each team (both default and custom) hands in **one** project proposal, which is worth 5% of your grade. This document specifies what information you should include in your proposal – it applies to both default and custom projects.

Before you start writing your project proposal, make sure that you have watched the *Practical Tips for Final Projects* lecture and read the custom project guidelines in section 4 of this document. If you're doing the default final project, you should also ensure you have read through the default final project handout (posted on the website).

We provide an example proposal template for your convenience:

https://www.overleaf.com/read/ccpkcdcscrzr

1 Key information to include

Your PDF proposal should have the following information at the top:

- **Title**: The title of your project (you can change this later).

- **Team member names**: List the names and @stanford.edu email addresses of all of your team members.

- **(Optional) External Collaborators**: If you have any collaborators who are not CS224n students, list them.

- **Custom or Default Project**: Indicate which you are doing.

- **(Custom project only) Mentor**: Note that we don't expect you to already have a mentor – you will be assigned one after we receive your proposal. Depending on your situation, write either (a) 'X has already agreed to be our mentor', (b) 'We would like to request X as our mentor', or (c) 'We have no particular mentor'. If you are doing one of the projects proposed by Stanford AI Lab members, write their name in case (a).

- **(Optional) Sharing Project**: If you are sharing this project between CS224n and another class, indicate it here.

2 Research paper summary (max 2 pages)

Paper selection: The first section of your proposal is a summary of a research paper that is relevant to your project. For example, if you are reimplementing a complex model, you should choose the paper that presents that model. If you are applying a particular method to a new task, you could choose the paper that presents the method. If you are developing a new variant of a particular model, you could choose the paper that presents the original model. There are other possible cases – use your judgment to choose what seems like the most relevant paper. If you are doing the default project, the papers linked in the default project handout are a great place to start.

Paper eligibility: Any research paper is fine, though you want to make sure it is a high-quality paper. For this, we encourage looking for papers that have been published in a peer-reviewed venue (i.e. conference or journal). However, due to the fast pace of Deep Learning and NLP research, there are many important and high quality papers that are not (or not yet) published at a peer-reviewed venue – for example, they may be very recent papers that have not yet been submitted for review. If you choose a paper that has not been published at a peer-reviewed venue, be wary that it might be unfinished or preliminary work.

The summary: Write a summary of the paper that a fellow CS224n student could understand. For most teams this will probably involve some key mathematical equations, but you don't have to exhaustively mathematically describe everything. You may include a diagram if you think it's important, but it shouldn't take up more than 1/3 of a page. In your 2-page (but not longer than 2-page) summary, prioritize conveying the most important information and ideas of the paper. In particular, when evaluating your summary, we will explicitly look for discussion of each of the following aspects:

Bibliographical info. Title, authors, publication venue (or preprint server, like arXiv), publication year, and URL. Put this information at the top of your summary.

Background. Set the scene for the paper, looking to the introduction section, as well as the related work or background sections, if they exist. What motivations and problems do the authors cite when explaining why they think this work is important? What problems are they attempting to solve, or what knowledge are they hoping to discover?

Summary of contributions. Each paper is published because it adds something to the ongoing research conversation. It teaches us something we didn't know before, or provides us with a tool we didn't have, etc. Summarize what contributions this paper makes, whether they be in new algorithms, new experimental results and analysis, new meta-analysis of old papers, new datasets, or otherwise.

Limitations and discussion. Every research paper has limitations and flaws. Using the discussion and conclusion sections if they exist, critically identify interesting experiments, methodology, or methods that might have made this paper stronger. For example, did the authors only evaluate on English, or only on Wikipedia text, and claim that their results generalize to all of language? Did the authors not characterize the errors their model makes compared to previous models? Discuss how these limitations contextualize the findings of the paper – do you still find the paper convincing?

Why this paper? There are infinite[1] papers you could read, and you chose to read this one. Maybe it came up first on Google Scholar, or a TA suggested it. . . regardless, discuss your motivation for choosing this paper or the topic that the paper it addressed. What interested you about the topic? Having read it in depth, do you feel like you've gained what you were hoping from reading it? ("No" is an okay answer here.)

Wider research context. Each research paper is a focused contribution, targeting a very specific problem setting. However, each paper also fits into the broader story of NLP research – designing systems that process human languages. In this course, we cover some fundamental concepts: how to represent language, what structure language has, why language is hard for computers to model, what problems tend to occur when applying deep learning methods to language. Connect the paper to these broad topics. Does the paper help us build better representations of language? If it helps us solve a particular task (like automatic translation or question answering,) do the methods have any promise for being more broadly applicable to other tasks (e.g., a new type of regularization in LSTMs applied in language modeling might be applicable to other NLP tasks!) It may be useful to do a cursory read of one or more of the papers cited in the paper you're reviewing, and cite them.

Do not go over two pages for this summary. Use your judgment to determine what information from the paper is important while making sure to devote a bit of space to each of the sections above. In particular, if you're doing the default project, you do not need to spend time explaining things that are already covered in the default project handout.

3 Project description (1-2 pages)

In this section, you will describe what you plan to do for your project. It's fine if your project eventually evolves into something different – that's a natural part of research. But your proposal should lay out a sensible initial plan. For custom projects, please read through Section 4 for guidelines and clarifications on topics and scope.

---

[1]Citation needed

This section should answer the following questions (it's a good idea to structure your project description in this way, but you can structure differently if you like):

1. *Describe the main goal(s) of your project.* If possible, try to phrase this in terms of a scientific question you are trying to answer – e.g., your goal may be to investigate whether a particular model or technique performs well at a certain task, or whether you can improve a particular model by adding some new variant, or (for theoretical/analytical projects), you might have some particular hypothesis that you seek to confirm or disprove. Your goal may be simply to successfully implement a complex neural model, and show that it performs well on a given task. Briefly motivate why you chose this goal – why do you think it is important, interesting, challenging and/or likely to succeed? If you have any secondary or stretch goals (i.e. things you will do if you have time), please also describe them. In this section, you should also make it clear how your project relates to your chosen paper.

2. *What NLP tasks(s) will you address?* Your project might be working with a pre-existing NLP task, in which case we would like a clear description of the task (i.e., give an example of an input and an output). Your project may also be addressing some topic in NLP, like an interpretability question, in which case we would like a clear description of how you plan to use a neural model to answer your research question: what inputs and outputs would you expect in your inquiry. The topics or tasks you describe here can be the same as those addressed by your chosen paper, but don't have to be.

3. *What data will you use?* Specify the dataset(s) you will use (including its size), and describe any preprocessing you plan to do. Be sure to specify whether the data is public and simple to download, and how you plan to access it if not. If you plan to collect your own data, describe how you will do that and how long you expect it to take.

4. *What neural method(s) are you planning to use?* Describe the models and/or techniques you plan to use. If it's already described in the paper summary, no need to repeat. If you plan to explore a variant to a published method, focus on describing how your method will be different. Make it clear which parts you plan to implement yourself, and which parts you will download from elsewhere. If there is any part of your planned method that is original, make it clear.

5. *What baseline(s) will you use?* How will you know that your method has worked better than a naive solution to the problem? Describe what methods you will use as baselines that you will compare your results to. Make it clear if these will be implemented by you, downloaded from elsewhere, or if you will just compare with previously published scores.

6. *How will you evaluate your results?* Specify at least one well-defined, numerical, automatic evaluation metric you will use for quantitative evaluation. What existing scores will you be comparing against for this metric? For example, if you're reimplementing or extending a method, state what score(s) the original method achieved; if you're applying an existing method to a new task, mention the state-of-the-art performance on the new task, and say something about how you expect your method to perform compared to other approaches. If you have any particular ideas about the qualitative evaluation you will do, you can describe that too.

Default final project teams will probably write shorter project descriptions than custom project teams. In particular, items (2), (3) and (6) are likely to be very short. There is no need to explain things that are already covered in the default project handout.

4 Guidelines and clarifications for custom projects

If you're doing a custom final project, you should feel free to choose a topic that interests you, and answer a question that you actually want to know the answer to — this is generally the way to stay motivated and do interesting research. However, since this is an NLP with Deep Learning class, we do restrict project topics to be within those areas. Our requirement is that all projects should be neural network NLP projects: your project must meaningfully interact with human language data (though this can be alongside other modalities), and utilize neural models.

Human language data Projects are expected to have at least some meaningful interaction with **human language data.** This means that projects that use deep learning methods discussed in lecture, including RNNs and Transformers, but do *not* interact with human language data, are out of scope for this class. This includes projects that exclusively work with code, biological/chemical data, or games (e.g. chess or go). However, note there are many ways of incorporating language into such projects which would make the project in-scope for this course. Projects cannot be on neural methods that are then applied only to non-language modalities like code.

Use of LLM APIs. Projects that partially, or even exclusively, make use of large language model (LM) APIs like OpenAI's GPT-3 or ChatGPT are **within scope** for a CS224n final project. However, students should also be aware that they are expected to make a substantive scientific and/or engineering contribution on top of such APIs. Since LM APIs abstract away a lot of the challenge of building NLP systems, the relative contribution of a project using an LM API is expected to be more than a project that works directly with LMs via deep learning libraries like Transformers or PyTorch.

With this in mind, projects that simply prompt an LLM like GPT-3 to generate text for a specific use case (e.g. summarizing news articles or generating song lyrics) are unlikely to be enough work for a final project. Examples of more appropriate contributions include (but are not limited to):

- Systematically identifying and/or benchmarking some capabilities, weaknesses, and/or biases of current LLMs

- Building substantial systems that interface with LLM APIs to enable new applications or workflows for end-users

Class project as part of a larger research project If you are doing your project in the context of an RAship or larger research effort, you should make it clear what part of the work you are doing **for this course**: though the overall research might span more than one quarter, what code/results do you start from for doing the final project for this course, and where do you hope to end up?

External collaborations For the final project, you are allowed to work with companies or universities outside of Stanford, though please keep in mind that the research cannot be so sensitive or private that you won't be able to produce a final poster and a final report for the course. Similarly, though you may work with private datasets, you will be expected to include an example or two of the type of data you use on the report or poster, so your data should not be so protected that this would be impossible. As mentioned before, you must make it clear in your proposal what part of the work of a longer research collaboration you are intending to do for this class.

5 Submission instructions

Submit your PDF proposal on Gradescope, and *make sure to tag all of your team members* – only tagged team members will receive credit. Additionally, when you tag the different 'questions' in your PDF, make sure to use the default project tags or the custom project tags, but not both.

6 What to expect

In grading the research summaries, we will be looking for well-considered analyses, and will be critically grading on the quality of the summary, not just based on completion.

We will grade your project proposals and provide brief feedback – for custom project teams, the grader will be the mentor. If there's a problem with your proposed project (e.g. it's not feasible in the given time), we may require you to submit a revised proposal – otherwise, your project is approved. All teams will be allocated Azure credits (based on the team listing in the proposal).

If you want to discuss your proposal before submitting it, you can go to the Office Hours of a knowledgeable staff member (look at the Office Hours webpage to see staff members' areas of expertise `http://web.stanford.edu/class/cs224n/office_hours.html`).