CS 224S
Final Project Guidelines and Suggestions

1 Project overview

One of CS224S' main goals is to prepare you to develop speech recognition and understanding systems of practical use. If you are interested in research, CS224S will also leave you well-qualified to do speech recognition and language understanding research. The class's final project will offer you an opportunity to do exactly this.

The important dates for the CS224S project are:

- Proposals: Due by the end of the day (11:59pm) on Mon, 5/1

- Milestone: Due by the end of the day (11:59pm) on Wed, 5/17

- Final poster: During the lecture time on Wed, 6/7.

- Final writeup (5 page limit): Due at noon on Thu, 6/8 (no late days).

Projects **must** be done in teams of three students. If you have a project which requires a larger or smaller team please email the course mailing list to request approval.

2 Project topics

Your first task is to pick a project topic. If you're looking for project ideas, please come to either Andrew Maas' or Jiwei Li's office hours, and we'd be happy to brainstorm and suggest some project ideas.

Many fantastic class projects come from students picking either an application or dataset that they're interested in and applying topics from the class to that task. Alternatively, if you're interested in specific set of techniques from the class we can help find a dataset or task where you can tractably explore those techniques.

If you're already working on a research project related to topics in class we encourage you to apply what you learned in class as a project. A very good CS224S project will comprise a publishable or nearly-publishable piece of work. In previous years, some number of students continue working on their projects after completing CS224S, and submit their work to a conference or journal.

For inspiration, you might also look at some recent spoken language understanding research papers. Topics covered in class span several conferences, but you can look at the recent proceedings of Interspeech, ASRU, SigDial, EMNLP, ACL, NAACL, NIPS, ICLR and ICML for research papers in this area.

Stanford has many datasets which you might use as a benchmark task for your project. You can browse the available datasets here:
`http://linguistics.stanford.edu/department-resources/corpora/`
Please contact the staff mailing list if you need help getting access to a dataset.

Titles of earlier CS224U projects are found here:
`http://web.stanford.edu/class/cs224s/earlyproject.html`

Projects will be evaluated based on:[1]

- The technical quality of the work. (I.e., Does the technical material make sense? Are the things tried reasonable? Are the proposed algorithms or applications clever and interesting? Do the authors convey novel insight about the problem and/or algorithms?)

- Significance. (Did the authors choose an interesting or a "real" problem to work on, or only a small "toy" problem? Is this work likely to be useful and/or have impact?)

- The novelty of the work, and the clarity of the writeup.

Lastly, a few words of advice: Many of the best class projects come from students working on topics that they're excited about. So, pick something that you can get excited and passionate about! Be brave rather than timid, and do feel free to propose ambitious things that you're excited about. Finally, if you're not sure what would or would not make a good project, please also feel strongly encouraged to either post on Piazza or come to office hours to talk about project ideas.

3 Project submission logistics

This section contains the detailed instructions for submitting different parts of your project. You probably do not need to read any of this in great detail until nearly the due date of the submissions.

3.1 Project proposals

Proposals are due by 11:59pm on Monday, 5/1. They should be submitted through Gradescope. A proposal should be up to 500 words and include the following:

1. The task you plan to work on.

2. The dataset you plan to use.

3. A sketch of your proposed approach/model.

4. How do you plan to evaluate your approach.

If your proposed project will be done jointly with a different class' project (with the consent of the other class' instructor), your proposal must clearly say so. The teaching staff will review your proposal and contact you if we foresee any issues with your project.

3.2 Milestone

Milestones are due by 11:59pm on Wed, 5/17. They should be submitted through Gradescope. A milestone report should include the following:

---

[1] Don't overthink these criteria, nor worry too much if you're not sure that you can do well on all of them. Just think of this as an "ideal" that you should aspire to (especially if your goal is to do publishable work).

1. Lit review of 2-5 relevant papers.

2. If your approach requires collecting data, describe your data collection set up.

3. Baseline results that you are comparing your model with.

3.3   Project presentations

Project presentation happens during the lecture time on Wed, 6/7. You will present your work in a poster session. Each group should prepare a poster, and be prepared to give a short oral presentation about the work. At the poster session you will also have a chance to see your classmates' projects. We will supply poster boards, pins, and easels for the poster session. The time and location of the poster session will be announced in class.

3.4   Final writeup

Final project writeups are due at 11:59am on Thursday, 6/8. They should be submitted through Gradescope **Late days cannot be used for the final writeup.** Final project writeups should be 6-8 pages long (including appendices, figures, references, and everything else you choose to submit). The following is a suggested structure used by CS224n for the final write-up:

1. Title, Author(s)

2. Abstract: It should not be more than 300 words

3. Introduction: this section introduces your problem, and the overall plan for approaching your problem

4. Background/Related Work: This section discusses relevant literature for your project

5. Approach: This section details the framework of your project. Be specific, which means you might want to include equations, figures, plots, etc

6. Experiments: This section begins with what kind of experiments you're doing, what kind of dataset(s) you're using, and what is the way you measure or evaluate your results. It then shows in details the results of your experiments. By details, we mean both quantitative evaluations (show numbers, figures, tables, etc) as well as qualitative results (show images, example results, etc).

7. Conclusion: What have you learned? Suggest future ideas.

8. References: This is absolutely necessary.

Please use ACL2017 template for writeups. The template can be downloaded here: `http://acl2017.org/downloads/acl17-latex.zip`

If you did this work in collaboration with someone else, or if someone else (such as another professor) had advised you on this work, your writeup must fully acknowledge their contributions.

After the class, we will post all the final writeups online so that you can read about each others' work. If you do not want your writeup to be posted online, please specifically mention it in the final paragraph of your writeup.