

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 on Friday, 5/2.
- Final presentations: In class Tuesday, 6/3.
- Final writeup (5 page limit): Due at noon on Monday, 6/9 (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 Prof. Jurafsky's or Andrew Maas' 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. If you are not already working on a research project but would like to, we have some suggested projects from PhD and post-doctoral researchers. Working on such projects can be a great way to get involved longer term in NLP and machine learning research.

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, and EMNLP for research papers in this area.

ASRU and Interspeech put their papers behind pay walls but you can usually find a copy somewhere by googling. Here are some recent links you can use:

<http://www.asru2013.org/accepted-regular-papers>
http://www.interspeech2013.org/Media/25-08-13_IS13_Program.pdf
<http://sigdial.org/sigdial-meetings>

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/departments-resources/corpora/>

Please contact the staff mailing list if you need help getting access to a dataset.

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

Project proposals should be emailed to the staff mailing list

cs224s-spr1314-staff@lists.stanford.edu

Proposals are due by 11:59pm on Friday, 5/2.

Your proposal should be a normal (plain ASCII) email, giving the title of the project, the full names of all of your team members, and about a 300-500 word description of what you plan to do. Please send your proposal as a normal email and not as an attachment, nor use any other document format (such as PDF or MS-Word). 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.

¹Don’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).

3.2 Project presentations

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.

3.3 Final writeup

Final project writeups are due at 11:59am on Monday, 6/9. **Late days cannot be used for the final writeup.** Final project writeups can be at most **5 pages** long (including appendices, figures, references, and everything else you choose to submit). Speech conference papers are often limited to 4-5 pages so by writing carefully you should be able to fit a clear presentation of your work.

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, there will be an option in the final submission script which you can use to indicate this.

Submission instructions: You'll submit a single PDF file for your whole group (not one copy per person) Make sure the PDF contains the project title and names of all your group members, including your SUNet email ids, on the first page.

Please submit your milestone in PDF format. In particular, we do not accept MS-Word, OpenOffice, PostScript, or any other document format. You should name your PDF file according to the format "`yourConcatenatedLastNames-ProjectTitle.pdf`". For example, if the project partners are John Doe and Jane Smith, and your project title is "Learning to Recognize People," then name your PDF file `DoeSmith-LearningToRecognizePeople.pdf`.

To submit the file, log into corn and create a submission folder, say `project`. Put your PDF file in it. Then in the terminal type the following:

```
/afs/ir/class/cs224s/bin/submit
```

and push enter. When prompted, fill up the required info like sunetID's and names. When it asks for the submission directory, just use `."` if you are inside your submission folder, or give it the path name for your submission folder.