

CS 229

Final Project Guidelines and Suggestions

1 Project overview

One of CS229's main goals is to prepare you to apply state-of-the-art machine learning algorithms to an application. If you are interested in research, CS229 will also leave you well-qualified to do machine learning or AI research. The class's final project will offer you an opportunity to do exactly this.

Keep in mind that this quarter, there **is no requirement** that you do a project—if you prefer to take a 24 hour take-home final, simply take the final; we will post logistics on this at a later date.

The important dates for the CS229 project are:

- Proposals: Due at 11 pm on Friday, 4/22 (late days allowed).
- Milestone: Due at 11 pm on Friday, 5/20 (late days allowed).
- Poster presentations: 12:15 - 3:15pm, Friday, 6/3/2016.
 - Poster is due as a PDF on Gradescope on 6/2 at 11 pm (no late days)
- Final writeup (5 page limit): Due at 11 pm on Monday, 6/6/2016 (no late days).

Projects can be done in teams of up to three students. If you have a project of such large scope and ambition that it cannot be done by a team of only three persons, you can propose doing a project in a team of four.

2 Project topics

Your first task is to pick a project topic. If you're looking for project ideas, please come to either Prof. Duchi's or the TAs' office hours, and we'd be happy to brainstorm and suggest some project ideas. In the meantime, here are some suggestions that might also help.

Most students do one of three kinds of projects:

1. **Application project.** This is by far the most common: Pick an application that interests you, and explore how best to apply learning algorithms to solve it.
2. **Algorithmic project.** Pick a problem or family of problems, and develop a new learning algorithm, or a novel variant of an existing algorithm, to solve it.

- Theoretical project.** Prove some interesting/non-trivial properties of a new or an existing learning algorithm. (This is often quite difficult, and so very few, if any, projects will try to do this.)

Some projects will also combine elements of applications and algorithms and theory.

Many fantastic class projects come from students picking either an application that they're interested in, or picking some sub-field of machine learning that they want to explore more, and working on that as their project. If you haven't worked on a research project before but would like to, you can also use this as an opportunity to try your hand at it. (Just be sure to ask us for help if you're uncertain how to best get started.)

Alternatively, if you're already working on a research project that machine learning might be applicable to, then working out how to apply learning to it will often make a very good project topic. Similarly, if you currently work in industry and have an application on which machine learning might help, that could also make a great project.

A very good CS229 project will comprise a publishable or nearly-publishable piece of work. Each year, some number of students continue working on their projects after completing CS229, and submit their work to a conference or journal.

So, for inspiration, you might also look at some recent machine learning research papers. Two of the main machine learning conferences are ICML and NIPS. You can also find papers from recent ICML conferences online:

<http://icml.cc>
http://icml.cc/2016/?page_id=175

All NIPS papers are online, at

<http://books.nips.cc/>.

Finally, to see a list of last year's class projects, you can go to

<http://cs229.stanford.edu/projects2015.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?)

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).

- 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 email us 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 at 11 pm on Friday 4/22 on Gradescope.

Your proposal should be a PDF document, giving the title of the project, the full names of all of your team members, the SUNetID of your team members, and about a 300-500 word description of what you plan to do. 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.

Submission instructions: Please upload your proposal to Gradescope under the "Project Proposal" assignment. It must be in PDF format. On Gradescope, only one team member must submit, but the person submitting must add all your team members to your submission on Gradescope. You can add team members to your assignment submission using your teammate's first and last name. If Gradescope cannot find them, make sure your teammate has registered on Gradescope.

3.2 Milestone

The project milestone is due at 11 pm on Friday, 5/20, which is roughly halfway between the proposal and the final project due dates. Your milestone report should describe what you've accomplished so far, and very briefly say what else you plan to do.

The milestone will help you make sure you're on-track. You should write it as if it's an "early draft" of what will turn into your final project. Specifically, you can write it as if you're writing the first few pages of your final project report, so that you can re-use most of the milestone text in your final report. Please write the milestone (and final report) keeping in mind that the intended audience is Prof. Duchi and the TAs. Thus, for example, you

should *not* spend two pages explaining what logistic regression is. Your milestone should include the full names of all your team members and state the full title of your project.

Submission instructions: Your milestone report should be at most 3 pages long. Please upload your milestone report to Gradescope under the “Project Milestone” assignment. It must be in PDF format. On Gradescope, only one team member must submit, but the person submitting must add all your team members to your submission on Gradescope. You can add team members to your assignment submission using your teammate’s first and last name. If Gradescope cannot find them, make sure your teammate has registered on Gradescope.

3.3 Poster presentations

The class projects will be presented at a poster presentation on Friday, 6/3/16. Each team should prepare a poster, and be prepared to give a very short explanation, in front of the poster, about their work. At the poster session, you’ll also have an opportunity to see what everyone else did for their projects. We will supply poster-boards and easels for displaying the posters.

3.4 Final writeup

Final project writeups are due at 11 pm on Monday, 6/6/16 to Gradescope. **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).

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.

Because the teaching staff will have only a few hours to see a large number of posters on 6/3, in the poster session we’ll only be able to get a gist of what you did. However, we (specifically, Prof. Duchi) personally read every word of every final writeup, and you can safely count on us getting the full details of what you did from the writeup. We know that most students work very hard on the final projects, and so we are extremely careful to give each writeup ample attention, and read and try very hard to understand everything you describe in it.

After the class, we will also 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, then please contact us at cs229-qa@cs.stanford.edu at least a week in advance of the final submission deadline.

Submission instructions: Please upload your final report to Gradescope under the “Final Report” assignment. It must be in PDF format. On Gradescope, only one team member must submit, but the person submitting must add all your team members to your submission on Gradescope. You can add team members to your assignment submission using your

teammate's first and last name. If Gradescope cannot find them, make sure your teammate has registered on Gradescope.

4 Miscellany

If, after CS229, you want to submit your work to a machine learning conference, the ICML deadline will probably be in early February next year, and the NIPS deadline is usually in early June (<http://www.nips.cc/>). Of course, depending on the topic of your project, other non-machine learning conferences may also be more appropriate.