CS129 Final Project Information

Project Proposal Due on Gradescope: Tuesday, Feb 3rd, 2026

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

Your project proposal should include the following information:

- Motivation: What problem are you tackling? Is this an application or a theoretical result?
- Method: What machine learning techniques are you planning to apply or improve upon?
- Intended experiments: What experiments are you planning to run? How do you plan to evaluate your machine learning algorithm?

Presenting pointers to one relevant dataset and one example of prior research on the topic are a valuable (optional) addition. If you are doing this project with another class, make sure to tell us which models you learned in this class you will be using.

Project milestone: Due on Gradescope: Friday, March 6th, 2026

The milestone will help you make sure you're on track, and should describe what you've accomplished so far, and very briefly say what else you plan to do. You should write it as if it's an "early draft" of what will turn into your final project. 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 Ng, Younes, 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. Note: We will expect your final writeup to be on the same topic as your milestone.

Contributions: Please include a section that describes what each team member worked on and contributed to the project. This is to make sure team members are carrying a fair share of the work for projects.

Grading: The milestone is mostly intended to get feedback from TAs to make sure you're making reasonable progress. As long as your milestone follows the instructions above and you seem to have tested any assumptions which might prevent your team from completing the project, you should do well on the milestone.

Format: Your milestone should be at most 3 pages, excluding references. Similar to to the proposal, it should include

- Motivation: What problem are you tackling, and what's the setting you're considering?
- Dataset: Describe the dataset you have now. How many training/validation/testing examples do you have? What kind of analysis did you do? What are the results?
- Method: What machine learning techniques have you tried and why?
- Preliminary experiments: Describe the experiments that you've run, the outcomes, and any error analysis that you've done. You should have tried at least one baseline.
- Next steps: Given your preliminary results, what are the next steps that you're considering?

Posters and Final writeups: Due on Gradescope March 17th, 2026

The class projects will be presented at a poster presentation. 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. You will also need to submit your poster as a PDF the day before the presentation. We will be using the same format as CS229. Here is a link to see the format we want poster guidelines

Final Write Up

- **Abstract (1 paragraph):** Abstract is optional, depending on your available space. It should consist of 1 paragraph consisting of the motivation for your paper and a high-level explanation of the methodology you used/results obtained.

- **Introduction (0.5 pages):** Explain the problem and why it is important. Discuss your motivation for pursuing this problem. Give some background if necessary. Clearly state what the input and output is. Be very explicit: "The input to our algorithm is an {image, amplitude, patient age, rainfall measurements, grayscale video, etc.}. We then use a {SVM, neural network, linear regression, etc.} to output a predicted {age, stock price, cancer type, music genre, etc.}." This is very important since different teams have different inputs/outputs spanning different application domains. Being explicit about this makes it easier for readers. If you are using your project for multiple classes, add a paragraph explaining which components of the project were used for each class.

- **Related Work (0.5 pages):** You should find existing papers, group them into categories based on their approaches, and discuss their strengths and weaknesses, as well as how they are similar to and differ from your work. In your opinion, which approaches were clever/good? What is the stateof-the-art? Do most people perform the task by hand? You should aim to have at least 3 references in the related work. Include previous attempts by others at your problem, previous technical methods, or previous learning algorithms. Google Scholar is very useful for this: https://scholar.google.com/ (you can click "cite" and it generates MLA, APA, BibTeX, etc.) Any citation format is fine.

- **Dataset: (0.5- 1 page)** Describe your dataset: how many training/validation/test examples do you have? Is there any preprocessing you did? What about normalization or data augmentation? What is the resolution of your images? Include a citation on where you obtained your dataset from. Depending on available space, show some examples from your dataset. You should also talk about the features you used.

- **Methods (1-1.5 pages):** Describe your learning algorithms, proposed algorithm(s), or theoretical proof(s). Make sure to include relevant mathematical notation. For example, you can briefly include the SVM optimization objective/formula or say what the softmax function is. It is okay to use formulas from the lecture notes. For each algorithm, give a short description (≈ 1 paragraph) of how it works. Again, we are looking for your understanding of how these machine learning algorithms work. Although the teaching staff probably know the

algorithms, future readers may not (reports will be posted on the class website).

- **Experiments/Results/Discussion (1-2 pages):** You should also give details about what (hyper)parameters you chose (e.g. why did you use X learning rate for gradient descent, what was your mini-batch size and why) and how you chose them. Did you do cross-validation, if so, how many folds? Before you list your results, make sure to list and explain what your primary metrics are: accuracy, precision, recall, etc... For classficiation projects, you include a confusion matrix. In addition, explain whether you think you have overfit to your training set, and what, if anything, you did to mitigate that. Make sure to discuss the figures/tables in your main text throughout this section. Your plots should include legends, axis labels, and have font sizes that are legible when printed. Make sure you state your errors, and accuracies.

- **Conclusion/Future Work (1-2 paragraphs):** Summarize your report and reiterate key points. Which algorithms were the highestperforming? Why do you think that some algorithms worked better than others? For future work, if you had more time, more team members, or more computational resources, what would you explore? Not all the team members have to be present in the presentation.

- **References and Contributions!**
  Final project writeups can be at most 5 pages long (including appendices and figures). We will allow for extra pages containing only references. 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 write-up must fully acknowledge their contributions. For shared projects, we also require that you submit the final report from the class you're sharing the project with.

  Feel free to adjust the specific sections according to your needs (e.g. combine introduction and related work or separate the experiments from the discussion. You are free to use single-column or two-column layouts. The paper size is standard A4 or 8.5 x 11 inches. Your font size must be greater than or equal to 10pt. Do not use less than 0.5 inch margins. You are not required to type your report in latex. If you use latex (or even Microsoft Word), we highly recommend using a conference/journal template (e.g. NIPS, IEEE, ICML). They generally provide both .tex and .doc templates. When you submit your final report, it must be in PDF format.