CIS 7000-005: Debugging Data & Models	Fall 2022

Project Checkpoint — November 3

Prof. Eric Wong

At this point there is approximately one month left in the semester. The goal of the checkpoint is twofold:

1. Help distribute the project workload over the semester

2. Provide a pivoting point in case things are not working out

With regards to the latter goal, at this point one of two things has likely happened.

The project is underway and preliminary results look good. In research, sometimes things go well. In this case, you should put together a concrete plan what you aim to accomplish by the end of the semester, based on what you have done so far. Most likely your results at this point are still prelimnary—you should think about what kind of experiments or results you would need to comprehensively demonstrate the result to a skeptical researcher.

The project has encountered some roadblocks or difficulties. In research, sometimes things don't go well. In this case, you should ask yourself: is there a different way in which you can achieve the original goal? If not, is there a different goal adjacent to your original project that is easier to achieve?

Structure of the checkpoint

You should also finalize most of the content from the proposal.

1. Proposal content - at this point, the content from the proposal should be mostly final (introduction, a comprehensive literature review in the related work, and a section describing the proposed work).

2. Current progress - Describe the work that you have currently done. What experiments or theories have you tried to run or prove? Were they successful or not, or did they end up being inconclusive? Most importantly, what can you actually conclude from these results?

3. Remaining work - Describe the work you would like to accomplish by the end of the semester. What kind of results or theory do you want to show? Importantly, why do you want to show them?

Some more detail on the last two sections below, as well as instructions for the checkpoint presentation.

1 Current progressl

When writing about results, here are some general guidelines on what you should and should not focus on. As a caveat, this is a very ML community driven way of writing, and the expectations for writing may bve different in other fields.

You should not do the following:

- Itemize a laundry list of experiments and their results. This is possibly the least interesting way to show experiments.

- Describe every last detail of experiments, such as training details, hyperparameters, etc. These belong in an appendix, not the main component of paper. Your average reader is not going to want to sift through this information.

- Full derivations for complicated theorems. This is too much unnecessary notation and information for your average reader.

Instead, you should do the following:

- Motivate each experiment/theorem. Why are you running this experiment or proving this theorem, and what question are you trying to answer, and whyu is it important? Connect it to the setting that you've (hopefully) described in your introduction.

- Give an intuitive, high-level description of what the experiment/theorem is.

- Interpret the result for the reader. Tell them exactly what they should conclulde from a given figure or table, instead of expecting them to infer the conclusion.

- Why is each result important? Connect the result back to the original motivation.

Remember, your target audience here is a general researcher in machine learning. Imagine your classmates, who have a wide variety of backgrounds and technical upbringings. You can think of this section as the "Results" section of a typical conference paper.

2 Remaining work

Here, you should outline next steps for your project. Keep the following in mind:

- Connect your next steps to the current progress. How does the current progress lead to or inform the remaining work?

- It is OK for remaining work to project further into the future than the remainder of this semester. That being said, it will be helpful to identify a subset of the work that you think is feasible by the end of the semester.

- Make the steps are concrete as possible. For example, specific experiments or theorems you wish you run or prove, and why you want to do them.

3 Format

Use the same format as the proposal. As a rough guideline you can aim for 3-4 pages.

4 Checkpoint presentation

Part of the checkpoint involves giving a 5 minute project presentation. The goal of this presentation is to let everyone else know the cool work you're doing, as well as practice distilling your research projects to the 5 minute format.

Distilling your research is a critical skill—you need to carefully think about what is the most important part for people to know, and how you can concisely convey this information. When you talk to people about your research, the vast majority will ask something like "What do you work on?" and you'll only have a few sentences to make an impression. Most people aren't going to sit around waiting for you to finish a 30 minute talk.

With that in mind you should focus primarily on conveying the follwoing items:

- Setting/motivation — Give a concise description of the setting or context of project that explains why your problem is important

- Problem — Give anm intuitive description of the problem you are solving and how you are doing so

- Current progress — What is the most interesting thing that you have discovered so far in this project?

You don't have to use these bullet points exactly for your slides. For example, it may make sense to combine the problem you're solving into the setting, or merge some progress into the problem solution. Use your discretion to decide what is the most concise and clear way to put this into a few slides with 5 minutes to spare. The objective is *clarity*, not complexity.

Slides. Prepared slides can be very powerful in effectively conveying information. However, you must be extremely critical of everythinig you put on a slide. Far too often, slides contain *way too much information.* The "text dump" slide is the worst offender here. These types of slides have the opposite intended effect—instead of helping you convey a message, they zone out the audience. Keep your slides simple and visual. If you have more than one sentence on a slide, you should have an extremely good reason to do so. One good way to calibrate is to ensure that each slide helps to convey exactly one message.

Groups If you are in a group, be sure that everyone has a chance to speak during the presentation (approximately) evenly. For 3 person groups, this may end up being 1 person per slide (i.e. setting, problem, current progress/future work).