Course Project CSE 599 — Advanced Machine Learning

The project for this course is designed to give you an opportunity (i) to engage with the current state of machine learning research and (ii) to contribute useful knowledge to this community. Your team will choose a direction from the list below and develop a project based on recent research:

1. Replication of recent work
2. Summarizing a line of theoretical work, for instance:
    - Summarize a line of work that aims to provide a theoretical explanation for an empirical phenomenon, e.g.:
        i. Neural networks generalize despite being a large hypothesis class
        ii. Learning of neural networks succeeds despite the loss function being non-convex
        iii. Mode connectivity in the loss landscape of neural networks
        iv. The “lottery ticket” hypothesis (related to initialization of neural networks)
        v. The lack of overfitting from test set re-use on ML benchmarks
    - Summarize a line of work that develop a new algorithm (e.g. extensions to SGD optimizers for large mini-batch sizes)
3. Original research, such as:
    - Proposing and evaluating a new idea on top of an existing code base
    - Testing the scaling properties of published methods that do not yet have scaling investigations
    - Analyzing the scaling properties of transformer variants or of specific datasets in conjunction with a regular transformer
    - Your own research project, if relevant

Teams: Your team should be composed of three people. We recommend that you include at least one CSE PhD student actively researching in machine learning or an aligned field. Beyond that we encourage you to have a mixed team with a CSE PhD student from some other research area, or someone who is not a CSE PhD student (different field or undergrad/masters student). Diverse teams are stronger teams; working with people whose perspectives are different from yours (both in that particular dimension and in other dimensions) gives you new opportunities to learn. You can use the team finding sheet to find a team or team members. If you are having trouble finding a team let us know (email dettmers@cs.washington.edu) and we will help you.

Your project will follow one of the directions listed above. We expect the following outcomes:

- Replication: Either failure (and why it did not work) or success in replicating a paper. Bonus points: an additional experiment that further analyzes the original authors’ work.
- Summarizing a line of theoretical work: Bringing together theoretical results from different papers / sources. Bonus points: Can you apply a theoretical framework to a

problem that is currently mostly empirical? Can you develop a new algorithm? Can you think of new experiments which test the main theoretical results?
- Original research: Implement something new and report your results. Bonus points: creativity, novelty, impact.

Some considerations in choosing a project direction:
- You should find the problem tackled in the research interesting.
- You should be able to access the data you will need to carry out the project.
- In many cases, the authors may have made code available; this may be a blessing or a curse. You should definitely evaluate a paper's codebase before deciding on that direction.
- You should estimate the computational requirements for the project and take into account the resources available to you. Some authors will have had access to infrastructure that is far beyond your budget; don't choose such a direction.

Deliverables and Deadlines

The project is split into three milestones. First, you submit a proposal which details what you want to work on. After that there is the first and final milestone. The first milestone is there to make sure you are on track. For replication or original empirical research this may mean that you found software that you can build on, and you tried running something on GPUs / make sure you have all the resources and requirements to begin with the experimentation.
For the theory projects, this may mean you surveyed the literature and mapped out how your main selected work relates to previous and follow-up work and how it relates to alternative perspectives.

Detailed instructions for the report are given in the Latex template provided on the website. It is imperative that you follow the instructions in the template carefully for versions 1 and the final version. Hence we strongly recommended that you familiarize yourself with the template before writing the proposal.

The deadlines for each deliverable are as follows:
- Proposal: Friday, April 28
- Version 1: Friday, May 12
- Final version: Friday, June 2

Proposal

Your proposal will be due around week 4. It should be one page and briefly include:
- a (bibtex) citation of the paper(s) or research direction you plan to base your project on, with a URL

- the goals and expected outcomes of your project
- a short description of whether and how you can access the data needed for the project
- whether you will use existing code (in that case, a link to the code) or implement yourself
- a discussion of the feasibility of the computation you will need to do (essentially, an argument that the project will be feasible)

Estimation does not have to be exact, but you will get more useful feedback if you include specific estimation. There is no specific template for the proposal.

Versions 1 and 2 (Final Report)

Fill out each section in the template by replacing the instructions with the actual content. You must follow this template for both versions 1 and 2 (final report). The final report must not exceed 8 pages, excluding references and the one-page summary that comes first. The appendix does not count towards the page limit.

For version 1 (due around week 6), you need only complete the following sections, and put a placeholder ("to do") for the other sections: Introduction, project scope, methodology.
For the final report (due at the end of the quarter), all sections in the template must be filled in. Please note:

- Grades are shared by your team. Students in this course are expected to work together professionally, overcoming the inevitable challenges that arise in the course of a team project. We recognize that, occasionally, team members behave unreasonably. Please get in touch with the TA Team to help resolve potential conflicts.
- No late submissions are allowed. Your team will receive zero points for the late submission.
- Instead of submitting code, set it up as a public Github repository and add the link to the project report.
- If writing your own code, make sure it is documented and easy to use (this project is about reproducibility!). Include a link to a github repository which can be installed and run with a few lines in bash on department machines. Include a description of how difficult the algorithms were to implement. If using public code from the original repository, more of your energy will go into running additional experiments, such as hyperparameter optimization, ablations, or evaluation on new datasets (see below). However, note that it's not always trivial to get a public code release working!
- You may include an appendix in your final report. However, you should include all the important details in the main paper. The appendix is allowed so that your report will be helpful to future researchers; it will not be graded by the course staff.

Grades
The project is worth 50% of your course grade, allocated as follows:

- The proposal is worth 10% of the project grade
- Version 1 is worth 25% of the project grade
- Version 2 is worth 65% of the project grade