Final Project

You are asked to carry out an original research project related to the course content. This description is intentionally very broad. Some possibilities include:

- Invent a new algorithm or architecture based on the ideas in the course.
- Try to explain a deep learning phenomenon that is poorly understood.
- Apply the techniques from the course to make a deep learning system work better. (This might be a good option if you're already using deep learning for your research, or are interested in doing so.)
- Many of the ideas in this course have only been tested in relatively simple settings like multilayer perceptrons or image classifiers. Find out if the conclusions still hold for other architectures, such as: RNNs, transformers, graph neural nets, GANs, autoregressive generative models, Bayesian neural nets, capsules, etc.
- Throughout the course, I sometimes speculate about why something happens, and note that it hasn't been tested. Rigorously testing these hypotheses could make a good project.

Collaboration. You should form teams of 2–3 students. The team does not have to be the same as for the Colab notebook assignment. Your final report should list the contributions of each team member. (See the FAQ below about individual projects and larger project groups.)

Project Proposal (due Feb. 17)

Each team will write a short (around 2 pages) research project proposal. It should include a description of a minimum viable project, some nice-to-haves if time allows, and a short review of related work. You don't have to do what your project proposal says — the point of the proposal is mainly to have a plan and to make it easy for us to give you feedback. (But if you want to completely change direction, it's a good idea to run it by the instructor first.)

Please submit your proposal through MarkUs.

Final Report and Source Code (due Apr. 7)

At the end of the class you'll hand in a project report, in the format of a machine learning conference paper (e.g. NeurIPS[1]). We recommend the report be about 8 pages plus references, but we do not enforce any minimum or maximum length.

Please submit your project report and the source code through MarkUs.

[1] `https://nips.cc/Conferences/2020/PaperInformation/StyleFiles`

Marking

85% of the marks will be given for meeting the requirements of the project and for the quality of the project proposal, presentation, and final report. (A mark of 85% corresponds to A-/A.) This includes:

1. **Abstract (5 points)** that summarizes the main idea of the project and its contributions.

   - Should be understandable to anyone in the course.
   - You don't need to say everything you did, just what the main idea was and one or two takeaways.

2. **Introduction (10 points)** that states the problem being solved and why we might want to solve it.

3. **Figure or diagram (10 points)** that shows the overall idea. The idea is to make your paper more accessible, especially to readers who are starting by skimming your paper.

   - For camera-ready diagrams, we recommend using Tikz, a LaTeX package. For Mac users, Omnigraffle is also convenient.

4. **Formal description (15 points)** of the model / algorithm / conjecture. Include at least one of:

   - An algorithm box.
   - Equations describing your model.
   - A theorem or formally stated conjecture.

   Highlight how your approach differs from existing work.

5. **Related work section and bibliography (15 points).**

   - If your project builds on previous work, clearly distinguish what they did from what your new contribution is.
   - Also, include a short (2–3 sentence) summary of other closely related papers.
   - I realize you might not know about all related papers (or have time to carefully read all related papers), and that's OK for this project.
   - Using BibTeX is annoying at first, but Google Scholar can give you the BibTeX entries.

6. **Comparison or demonstration (20 points).** Include at least one of:

   - A proof of a theorem or conjecture, or an interesting counterexample.
   - An experimental comparison of the results of your method compared with a baseline. Qualitative evaluation is OK.
   - An experiment demonstrating a property that your model has that a baseline model does not. Experiments should also include a description of how you prepared your datasets, how you trained your model, and any tricks you used to get it to work.

- An experiment that reveals interesting properties of or relationships between existing methods.

Toy data is OK! The point is to help the reader understand why or when we would want to use one approach over another, or to understand something better. Try to summarize the main takeaways. Negative results are fine, as long as you have an insightful and well supported explanation.

7. **Limitations (5 points)** of your approach.

   - Describe some limitations of your proposed method.
   - Give some examples of possible extensions, ways to address these limitations, or open problems.

8. **Conclusions (5 points)**

   - State the results achieved in relation to the problem described in the introduction.
   - Repeat the main takeaways from your paper.

9. **List of contributions.** You should list the contributions of each team member.

The other 15% will be given for going above and beyond. We are looking for some form of creativity, e.g. clever experiments that reveal an interesting phenomenon, tricks for circumventing obstacles, etc.

Standards will be higher if you are doing a project related to your research.

FAQs

Can I do a project that overlaps with my research? You are welcome to do a project related to your research. Your project proposal and final report must **each** clearly explain the relationship to your research, what work was already done prior to the course, and what work (if any) was done by people not on the project team. Our expectations will be higher compared with completely de novo projects.

Can I do a project that overlaps with a project for a different class? No.

Can I do the project solo? You are allowed to do the project solo if you really want to. However, the standards will be just as high as for group projects. Also, your project report will be graded by a TA rather than the instructor.

Can we form a group larger than 3? Groups larger than 3 require permission of the instructor, and the standards will be considerably higher.

Do I need to use JAX? You are welcome to use whatever language and framework you like for the project. Be aware that some important operations for the course, such as directional derivatives or Hessian-vector products, might be much harder to implement in some frameworks.

What if I find related work? If you discover partway through your project that much or all of the work you did has already been done, that's OK. Just make a note of this in your report. It's unreasonable to expect you to have mastered the literature before starting the project.

What if it doesn't work better than the baseline? That's OK. Not all ideas work, which is why we need to do research. What matters for your grade is the quality of your analysis. You should consider various hypotheses for why your method does or does not work, and figure out how to tease them apart experimentally. What did we learn that we didn't know before you ran the experiments? (Of course, this goes for positive results as well as for negative ones.)