Project Report Requirements

Machine Learning: 10-707

Spring 2021

1 Course Project Guidelines

Your class project is an opportunity to explore an open problem in deep learning, that is either of a mathematical, scientific or engineering nature.

- If you are doing a *mathematical project*, it should at least include a detailed survey of recent works on the topic; it can include novel theoretical results or conjectures corroborated by synthetic and real-world data experiments.
- If you are doing a *scientific project*, it should include well thought-out experiments, documenting and exploring some phenomenon in deep learning. It should include both synthetic and real-world data experiments, as well as some hypotheses generated by the experiments. These may or may not be accompanied by theoretical results.
- If you are doing an *engineering project*, it should involve picking an interesting dataset and task, applying one or more basic deep learning approaches to establish a performance baseline, then going beyond your baseline approach to develop and study more sophisticated and hopefully more successful approaches.

You should work in groups of two (2) students. Exceptions may be granted on a case-by-case basis for larger or smaller groups. Team members are responsible for dividing up the work equally and making sure each member contributes.

Your project will be worth 30% of your final class grade, and will have two deliverables:

1. Midway Report, due Mar 29: 3 pages (not including references) (20%)
2. Final Report, due May 7: 8 pages (not including references), plus unlimited appendix for figures + proofs. (80%)

2 Grading Criterion

Each deliverable of your project will be evaluated based on several factors:

1. Creativity: The groups are encouraged to come up with original ideas and novel applications. A project exploring new ideas (algorithms, methods, theory) on ML or new, interesting applications of existing algorithms is scored higher than a project without many new ideas/applications. Some useful tips about evaluating the quality of the contributions in a paper here: `https://2020.emnlp.org/blog/2020-05-17-write-good-reviews`.
2. Completeness: The extensiveness of the study, experiments, and analysis of results. A project that produces a more intelligent system by combining several ML techniques together, or a project that involves well-designed experiments and thorough analysis of the experimental results, or a project that nicely incorporates various real world applications, are scored higher.

3. Clarity of writing: The report should be organized clearly and well written.

4. NIPS format: Use NIPS format for all your reports. Length: Don't exceed the page limit. For additional figures/proofs, use the appendices. You should assume the main 8 pages will be read carefully, appendices will be skimmed.

5. Structure: See the following section for more detail.

3 Report Structure

3.1 Midway Report

1. Title

2. Introduction: Which project did you choose? What are you trying to solve? You should include a formal description/background about the problem and data/experiments that you chose, and how you plan to analyze your results.

3. Background/Literature: You should research work related to the problem. What problem did they solve and how does it relate to yours? How can you improve on what has already been done?

4. Methods/Model: By the proposal, if you are doing a project with an empirical component (either synthetic or real-world data), you are expected to have implemented some baseline methods or small-scale experiments. If you're doing a mathematical project, you should have a summary of the techniques in prior works, as well as your proposed line of attack for future results.

5. Preliminary Results: Show plots of the performance of your baseline algorithm and/or preliminary theorems/conjectures.

6. Evaluation of preliminary work: Evaluation of your baseline method and any further algorithms that you have tried. How do you hope to improve on the work which you have already done? What was successful and what was unsuccessful?

7. Future work: Which techniques do you plan to apply to beat the baseline method / prove new results ? What is your motivation behind these techniques (you are highly encouraged to come up with an original idea of your own or interesting applications rather than simply implementing or applying existing ML algorithms or doing a literature survey)? How do you plan to evaluate your final method? Goals, timeline, and division of work throughout your team. Provide a rough timeline of your plan ahead and job to be done by each team member.

8. Teammates and work division: We expect projects done in a group to be more substantial than projects done individually. You should outline what everybody in your group will do and by when each task should be complete.

9. References and citations: Clean and correctly formatted citations and bibliography.

3.2 Final Report

1. Title

2. Introduction: What are you trying to solve? Why is it important? What does your data look like?

3. Background: Briefly summarize the findings from your midway report. Please do not include any "boilerplate" content (e.g., descriptions of the domain, standard background on Deep Learning, explanations of CNNs or other methods implemented in your homeworks).

4. Related work: Previous work related to your topic that you may have referenced to help guide your project.

5. Overview of results: Describe your results: what new method did you provide? What theorems did you prove? What hypotheses did you generate and what conclusions did you draw? What metrics did you use for evaluation? How do your results compare to prior work?

6. Methods/experimental setup/proofs: Describe in detail your approach. What is your method? How did you prove your theorems? What is the experimental setup and what did the results show?

7. Discussion and analysis: Analyze your model and results. Highlight a few limitations of your approach (e.g., strong assumptions you had to make, constraints, when your method did not work in practice, etc.). Do the results and the explanation provide insights into the ML models or the environment that you were dealing with? Comment on whether you think there is a way to further improve your method to eliminate these limitations.

8. References and citations: Clean and correctly formatted citations and bibliography.