# Project Report Requirements

Machine Learning: 10-417/10-617

Fall 2022

## 1 Course Project Guidelines

Your class project is an opportunity to explore an interesting machine learning problem of your choice in the context of a real-world data set. You should propose a project based on your chosen dataset, and provide us a pointer to that data when submitting your midway report.

A typical project consists of (1) picking an interesting dataset and task , (2) applying one or more basic deep learning approaches to establish a performance baseline , then (3) going beyond your baseline approach to develop and study more sophisticated and hopefully more successful approaches. The number of approaches you explore in (3) will probably depend on the number of members in your team.

1. Projects can be done in teams of 2-3 students. Team members are responsible for dividing up the work equally and making sure that each member contributes.

2. Can a single student do a project on their own? Only in the following situation: (1) you want to design a project around a data set associated with your ongoing research, and (2) you absolutely cannot find another student in our class who would be interested in your problem, and (3) your faculty advisor associated with that ongoing research writes us an email stating that she will be happy to serve as mentor for your class project. (Your advisor should email Brynn with the subject "10417/10-617 Project Approval" bedmunds@andrew.cmu.edu)

3. Each project has a TA as a project consultant/mentor.

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

1. Midway Report, due Oct 31: 3 pages (not including references) (5%)

2. Final Report, due Dec 07: 8 pages (not including references) (25%)

## Grading Criterion

The page limits are strict! Papers over the limit will not be considered. 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 for working in their environments. 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.

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. NeurIPS format: Use NeurIPS format for all your reports. Length: Don't exceed the page limit.

5. Structure: See the following section for more details

## Report Structure

### 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 that you chose. It should be clear what your inputs and outputs are, and how you plan to analyze your results.

3. Background/Literature: You should research work related to your own. 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, you are expected to have implemented a baseline (i.e., a standard deep learning approach to which you will compare your method). You should discuss what method(s) you have implemented and show plots of the performance. Additionally, if you have already gone beyond the baseline, you should describe what work you have completed towards creating a method which beats the baseline.

5. Preliminary Results: Your experimental results. Show plots of the performance of your baseline algorithm and interpret what they mean. Be sure to label and explain this clearly.

6. Evaluation of premilinary 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? 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)? 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.

**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. Methods: By the final report, we expect you to have implemented your own ideas beyond the baseline. Additionally, you should describe what work you have completed towards creating a method which beats the baseline. In addition to successful approaches, you should briefly detail approaches which you tried and found to not work well. What methods have you completed? 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)?

6. Results: Your experimental results. Show plots of the performance of your algorithms and interpret what they mean. Be sure to label and explain this clearly. Describe how the current results in each of the experiments

align with your expectations. What metrics did you use for evaluation? How do your results compare to prior work?

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.