10-608: Course project guidelines*

Igor Labutov, Bishan Yang

January 2018

1 Overview and Timeline

The project is intended to be a investigation of your choice in the topic areas relevant to the class (i.e. "machine learning from human instruction", "instruction and question answering over personal knowledge", "machine learning from natural language feedback"). The purpose of the project is to do research in an area of your interest in greater detail. It should be scoped to be original and publishable research. The project can take on many forms, including but not limited to:

- Projects that explore the application of machine learning ideas to an interesting "real world" problem.

- Projects that involve a theoretical or empirical study of aspects of a learning method or model.

- Projects that extend existing learning algorithms, or that propose new learning algorithms.

Please propose a topic to us in your project proposal, and we will give you feedback on its feasibility. You will present the results of your project both in a short presentation in class, as well as in a final project report. More details are given below.
The important dates are:

- **February 7:** Form groups (no submission due). Groups of at most 2 people are allowed

- **February 11** Project proposals due

- **February 13** Project proposal presentations

- **March 20** Midterm reports due

---

*this document is based on the project structure document created by Thorsten Joachims at Cornell for the Advanced Topics in Machine Learning course

- **April 24** Final project presentations

- **May 1** Poster presentations

- **May 10** Final project reports due

2   Project proposal

The project proposal should outline what you want to do in your project. We and your peers will give you feedback on your proposal to make sure the project you are proposing is feasible and appropriate. The proposal should contain the following sections:

- **The Team**. Names of the people working on the project.

- **Motivation**. Explain why this project is interesting and important.

- **Statement of the Problem/Task**. A statement of the problem, issue, or task that you're interested in studying. In particular, try to formulate the key questions (2 to 4 questions is probably a good number) that you will answer in the project.

- **General Approach**. A high-level description of the general approach you'll use to address the questions. This should include how you will evaluate and what evidence you are planning to gather (e.g. how you can answer the questions through experiments on data).

- **Baselines**. Baseline models that you foresee comparing against

- **Resources**. A list of resources you have/need to conduct the project. This includes reading, software, datasets, etc. How are you planning to get these resources?

- **Schedule**. A schedule of work indicating the dates by which you plan to complete components of the project. Make sure the schedule is plausible.

The maximum length for the proposal is two pages in ACL style (including title, authors, and references). The ACL style is available at http://acl2018.org/call-for-papers/.

3   Status Report

The status report contain a refined version of the proposal after incorporating instructor feedback and the peer feedback, and the description of the work that you have done up to that point. The report should contain your progress, any preliminary results, and as well as a summary of the changes that you made to the project proposal (total length 4 pages).

4   Project Presentation

The project presentation will be during the last two weeks of class. We will assign a time slot to each group. In the presentation, you should

- provide motivation for your project, explaining why it is important and interesting,

- explain your research questions,

- provide evidence,

- draw conclusion.

You can use the computer projector (e.g. via powerpoint) for your talk. Try to make the presentation interesting (e.g. by including a demo). More details on the logistics follow later.

5   Project Report

The final project should be submitted via CMT as a PDF file. There is no need to submit any other material like raw data, experiment logs, or source code. A good model for writing the report is imagining that you are writing a conference paper. You have seen many good conference papers during the semester, so you should have an idea what a good paper looks like. The most important points to keep in mind for the report are

- to motivate your project (if you are using an already defined project, then to motivate your approach to the problem)

- to discuss related literature, and to differentiate your work

- to describe your approach, model in detail, providing motivation and reasoning behind your approach

- to be **thorough** in thinking through various baselines, either simpler versions of your model, or competing models.

- to make a convincing argument that supports your conclusions; this means having conducted experiments that support your hypothesis

- to make sure that credit is given to all software, literature, etc. that helped you in your work.

Also, keep in mind the criteria according to which the project will be evaluated (see section "Review Criteria and Grading"). The length limit for 8 pages (same as an ACL paper), and most research results can be fit into that length. Being concise is a good thing, but do not sacrifice clarity and completeness.

6 Peer Reviewing

You will receive and give peer feedback on all stages of the project, namely the proposal, the status report, the presentation, and the report. These reviews will be single blind, which means that the reviewer knows the author identities, but the author does not know the reviewer identities. Part of the code of academic integrity is to not review authors you feel conflicted with, and to not reveal to any of your peers which papers you reviewed. When you write reviews, the goal is to be as objective and constructive as possible. Also, please be polite!

7 Review Criteria and Grading

The projects will be graded in the same spirit as research papers are assessed. Here is a list of things that we will be looking for:

- Originality (how novel is the research)

- Significance (are the questions you are asking interesting and important)

- Relevance (is the work relevant to the course)

- Quality of arguments (are claims supported, how convincing are the arguments you bring forward)

- Clarity (how clearly are goals and achievements presented)

- Connections to related work

- Scope/Size (in proportion to size of group)

Relative to each other, the proposal will account for 10% of the grade, the status report 10%, the presentation for 20%, and the report for 60%. Feel free to come and talk to me during my office hour about the various aspects of your project so that we can make sure that you are on the right track.

Have fun, and good luck!