CMPSCI 689: Machine Learning
Final project guidelines

Final projects: Final projects will entail original investigation into any area of machine learning defined very broadly. That means that machine learning over visual data, or text, or speech, language-vision interfaces, and so on, are all acceptable topics, in addition to the core machine learning topics.

Scope: As a broad target, the final projects should involve approximately as much work as ~~two~~ one homework assignment. For groups, the total work should scale roughly linearly with the group size, and be distributed roughly equally. Similarly, multi-purpose projects which are being submitted for multiple classes should scale with the number of classes involved. An ambitious, well-done project from a group of two or more (or shared between two or more classes) should be on the order of a conference paper in depth of experimentation. I encourage you to tackle large problems in groups, for multiple classes, or both.

Team size: Each project must be done in a group of size two or more.

Grading and Milestones: The milestones will be:

April 2nd	Abstract due
April 23rd, 27th	Project presentations
May ~~3rd~~ 6th	Final reports due

Abstract. The abstract is just a short paragraph telling me who is in your group, describing the problem you've chosen, sketching the general approach you intend to take, and stating the kinds of data you're using. If you haven't already spoken to me about project ideas, you may want to stop by my office hours or to make an appointment before this point. The abstract mainly serves to give me a chance to make sure you're on a good path and to help me get a sense of who is doing what. Hard copies of the abstracts must be submitted in class.

Presentation. Each team will make a short presentation (8 mins including questions) describing their preliminary results. An important skill in research is to be able to tell in a week or two whether your ideas are basically going to work, well before you've fully done all engineering and experiments. The order of the presentation will be randomized so be prepared to present your work on either April 23rd or 27th.

Writeup. The final write-up should be on the order of 6-8 pages in the NIPS paper format (style files can be obtained from – `https://nips.cc/Conferences/2014/PaperInformation/StyleFiles`), describing your approach, results, data analysis, and so on. The initial abstract is a required checkpoint, but you will only receive a grade at the end, based on your final write-ups. Under normal circumstances, all group members will receive the same grade for the final project. Late days will not apply to the final reports. I have to get your grades in to the university, and I'm already giving you as long as I possibly can. The final report should be uploaded as a pdf to Moodle.

Ideas: You are welcome to come up with your own topics – some of you already may have done so. Take a look at the the resources listed below for potential topics. You are also welcome to come by my office hours to get ideas from me.

- Carlos Guestrin's course at CMU: `http://www.cs.cmu.edu/~guestrin/Class/10701/projects.html`
- Another course at CMU: `http://www.cs.cmu.edu/~ggordon/10601/projects.html`
- Andreas Krause's course at Caltech `http://courses.cms.caltech.edu/cs253/projects.html`

In addition Kaggle (`https://www.kaggle.com/competitions`) has a bunch of competitions that you can participate.