Instructions for Course Project

1 Project Proposal

Please submit by Thursday 11/1 if possible but not later than Monday 11/5. Submission is by email to roni@cs.tufts.edu.

Please write two paragraphs per project proposal.

The first paragraph describes your data or application and the regression/classification problem associated with it. The second paragraph describes what probabilistic model or algorithm may be used in this context.

I would like to see that the project is not "too hard", that is, one can complete it within the allocated 4-5 weeks. I would also like to see that the project is not "too easy" meaning it requires a good amount of work and some investigation.

A project can be "application focused" meaning a major component of your work is devoted to collecting and representing the data, and different choices will be investigated within the project. A project can also be "methods focused", meaning that a major part of the work is devoted to comparing, developing and evaluating some less well understood aspects machine learning methods.

I also put some project ideas, which you can consider, on the course web page. The first two might be easier than others so you can consider these to be "default" project topics. If you have any questions, or want to discuss some ideas for your project, please email roni@cs.tufts.edu and we can either discuss by email or set up a time to meet. Please do so before the deadline so we can get you started on the project in good time.

2 Project Submission

Final project reports, hardcopy only, should be submitted by Monday 12/10 5pm in Prof. Khardon’s mailbox in the CS main office.

Please investigate the application, machine learning methods, algorithms, and so on as agreed in the proposal and run an experimental evaluation to the best of your ability/knowledge. Then write a report on the project. The best way to think about the report is as a paper submitted to a machine learning conference (ICML, ECML, NIPS, ...) or the relevant conference in the application domain. In any case the report should include:

(1) Description of problem/application, data representation and goal for machine learning in that context.
(2) Description of machine learning methods with details for the parts that are not straight from
the textbook or lectures and/or (2') Description of the machine learning aspects w.r.t. methods that are being investigated.

(3) Description of the experimental setup and how it addresses issues from (1) or (2'). NB if this is an "application project" then issues are in (1). If this is a "methods project" then issues are in (2'). Most likely you have issues to discuss in both.

(4) Experimental results and discussion of what they tell us regarding issues from (1) or (2').