Statistical Machine Learning (600.675 Spring’ 2015)

This is a second graduate level course in machine learning. It will provide a formal and an in-depth coverage of topics at the interface of statistical theory and computational sciences. We will revisit popular machine learning algorithms and understand their performance in terms of the size of the data (sample complexity), memory needed (space complexity), as well as the overall computational runtime (computation or iteration complexity). We will cover topics including nonparametric methods, kernel methods, online learning and reinforcement learning, as well as introduce students to current topics in large-scale machine-learning and randomized projections. Topics will vary from year-to-year but the general focus would be on combining methodology with theoretical and computational foundations.

Schedule
Lectures
    Thursdays, 12PM-2:30PM in Malone 228 (2/19 lecture in Malone 107)
Office hours
    Raman Arora: Thursdays 2:30PM-3:30PM in Malone 228
    Poorya Mianjy: Tuesdays 9:30AM-10:30AM in Malone 338
Recitation
    Poorya Mianjy: Fridays 9:30AM-10:30AM in Malone 328

Contact Information
Instructor: Raman Arora  Malone 331  arora@cs.jhu.edu  516-1327
TA: Poorya Mianjy  Malone 338  mianjy@jhu.edu

Prerequisites
You should have taken CS 600.475 or CS 600.476 or equivalent introductory machine learning class. We will assume that you are familiar with the following concepts:

1. Linear algebra (vector spaces, orthogonality, singular value decomposition)
2. Probability and Statistics (random variables, probability distributions, expectation, mean, variance, covariance, conditional probability, law of large numbers, Bayes rule, MLE)
3. Introductory machine learning (classification, regression, empirical risk minimization, regularization)
4. Convex optimization (differentiation, unconstrained optimization, constrained optimization, Lagrangian, KKT conditions)

We will hand out notes on these topics and cover them briefly in the first lecture and in recitations but it is your responsibility to make sure you understand these concepts.
Text
We will use the following textbooks for this course:

Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. 2014. Q325.5 .S475 2014 c. 1

Other useful references are:

Grading
The coursework for CS.675 includes:
1. **Four assignments.** They are due on **Thursdays at 12PM** before the class. Hand them to Poorya Mianjy (Malone 338). If he is not available, write your name and date and time of submission and slide them under my office door (Malone 331).
2. **Midterm exam.** The exam will be on April 9th.
3. **Project.** The project involves writing a short paper summarizing key results in literature on a particular topic. You will pick a topic of interest, read relevant papers in the area and present a summary of theoretical analysis along with proof sketches. You may work by yourself or in teams of two.

A one page proposal will be due on **Thursday, February 19th**. Your proposal should include (a) title and list of team members, (b) a clear problem description and scope of the project, and (c) reading list.

A brief progress report will be due on **Tuesday, March 31st**. The progress report should include (a) a strong well-motivated introduction, (b) a summary of work you have done so far and (c) what remains to be accomplished.

The final report (maximum 8 pages) will be due on **Thursday, April 23rd**. The final report should be in NIPS format (latex style files here: http://nips.cc/Conferences/2014/PaperInformation/StyleFiles) and include abstract, introduction, notation and assumptions, key results and conclusion section commenting on the meaning of the results and open questions. The reports should be self-explanatory. If you are working
on an application oriented project your final report should include 
abstract, introduction, related work, methodology, theoretical/experimental results, discussion 
of results and conclusion section.

Grading will be based on homework assignments (30%), final project (40%), in-class 
midterm exam (30%).

**Topics**

A tentative list of topics that we will cover:

1. Intro [Motivation, definitions, terminology, probability tools, concentration 
   inequalities]
2. Foundations [PAC learning, finite class, realizable case, unrealizable case, Bayes error, 
estimation and approximation errors, model selection, regularization, Rademacher 
   complexity, VC dimension]
3. Nonparametric methods [density estimation, nonparametric Bayes, clustering]
4. Support Vector Machines [linear classification, margin bounds, kernel methods]
5. Ensemble methods [Weak and strong learning, margin interpretation, Adaboost, 
analysis]
6. Online learning [Perceptron, online gradient descent, experts setting, winnow rule, 
   Bregman divergence, mirror descent, online to batch conversion, sketching]
7. Regression [Generalization bounds, linear regression, kernel ridge regression, support 
   vector regression, sparsity]
8. Large-scale machine learning [Big Data, parallel and distributed learning, mini- 
batching, stochastic approximation algorithms, SGD methods, Pegasos]
9. Other topics [Active Learning, Reinforcement learning, Ranking, Randomized 
   projections]

**Schedule**
The class schedule is available [here](#) and will be updated throughout the semester.