

CS 281A / Stat 241A Fall 2009

Statistical Learning Theory

Web site: <http://www.cs.berkeley.edu/~bartlett/courses/2009fall-cs281a>
Please check the web site for announcements.

Lectures: Soda 306. Tuesday/Thursday 11-12:30.

Instructor:

Peter Bartlett (bartlett@cs). Office Hours: Tuesday 3-4pm, 399 Evans Hall.
Wednesday 3-4pm, 723 Sutardja Dai Hall.

GSIs:

Alekh Agarwal (alekh@cs) [Office hours: TBA]
Joe Neeman (neeman@stat) [Office hours: TBA]

Discussion Section: Time and place to be determined.

Course description

This course will provide an introduction to probabilistic and computational methods for the statistical modeling of complex, multivariate data. It will concentrate on graphical models, a flexible and powerful approach to capturing statistical dependencies in complex, multivariate data. In particular, the course will focus on the key theoretical and methodological issues of representation, estimation, and inference.

- Introduction to graphical models.
- Conditional independence. Directed and undirected graphical models.
- Inference:
 - Elimination algorithm.
 - Sum-product algorithm.
 - Factor graphs.
- Estimation and parameterization:
 - Bayesian, maximum likelihood, MAP estimation.
 - Linear regression.
 - Linear classification.

- Exponential family. Conjugacy. Sufficient statistics.
- The EM algorithm.
- Examples:
 - Hidden Markov models.
 - Factor analysis.
 - Kalman filter and smoother.
- General inference algorithms:
 - Junction tree.
 - Approximate inference: sampling methods
 - Approximate inference: variational methods

Prerequisites: Previous coursework in linear algebra, multivariate calculus, basic probability, statistics, and algorithms. Previous coursework in graph theory, information theory, and optimization theory would be helpful but is not required.

Familiarity with a matrix-oriented programming language (such as numpy, R, Splus, or Matlab) will be necessary.

Textbook: *An Introduction to Probabilistic Graphical Models*, by Michael I. Jordan. Available as a reader from Copy Central at 2483 Hearst Ave.

Assignments: The grade will consist of 60% for regular homework assignments (approximately one every two weeks) and 40% for a substantial project. It is fine to discuss homework assignments with other students, but all students must write up their own individual homework assignments and indicate with whom they have discussed the homework problems. A late homework will have its grade reduced by $\lceil t_{\text{late}} \rceil \times 33\frac{1}{3}\%$, where t_{late} is the number of calendar days that the homework is late. The project will involve independent work on a topic related to the course.

- Friday, October 30: brief project proposal due (one or two paragraphs).
- Monday, November 30: final project report due.
- Week of November 30: poster presentations of the projects.

Academic Dishonesty: Please see the policy on academic dishonesty at <http://www.eecs.berkeley.edu/Policies/acad.dis.shtml>.