

Statistics 231 / CS229T: Statistical Learning Theory

John Duchi, Stanford University, Spring 2017

Prerequisites

There are no formal prerequisites to this class. But we will assume a significant level of mathematical maturity. This means an understanding of the following.

- Machine learning: at least at the level of [CS229](#)
- Linear algebra: a working knowledge at the level of [EE263](#) or [Math 104](#)
- Probability: this course will have substantial probabilistic content and require non-trivial command of probabilistic techniques. The absolute bare minimum is probability at the level of [Stats116](#)
- Convex optimization will be extremely helpful, but is not strictly necessary. Simultaneously taking the course [EE364a](#) is recommended because you will find its content extraordinarily useful more broadly

Texts

There is no required text for the course, though we will post lecture notes for each of the lectures. There are a number of useful other references. Feel free to contact us with more if you find useful ones.

- [Percy Liang's course notes from previous offerings of this course](#)
- [Peter Bartlett's statistical learning theory course](#)
- [Boyd and Vandenberghe's Convex Optimization](#)
- [Sham Kakade's statistical learning theory course](#)
- [Martin Wainwright's statistical learning theory course](#)

Grading

Your grade will be determined by homework (40%), two paper reviews you will write (20%), and a final project (40%).

- Homework: There will be (approximately) one homework assignment every two to three weeks throughout the course, which will count for 40% of the grade. In effort to speed grading and homework return to you, we will grade homework problems and their sub-parts on a $\{0, 1, 2\}$ -scale: 0 indicates a completely incorrect answer, 1 indicating approximately halfway correct, 2 indicating more or less correct.
- Paper reviews: we will post more information shortly, but you will (at some point in the quarter) be required to read and write a short review, of about 4 pages, of two papers.
- The final project will be on a topic plausibly related to the theory of machine learning, statistics, or optimization.

Course Overview

When do machine learning algorithms work and why? This course focuses on developing a theoretical understanding of the statistical properties of learning algorithms.

Potential Topics

- Concentration inequalities: sub-Gaussian and sub-exponential random variables, randomized embeddings
- Uniform convergence: uniform laws of large numbers, VC dimension, entropy integrals, Rademacher complexity
- Online learning: online convex optimization, mirror descent, multi-armed bandits, concentration
- Kernel methods: reproducing kernel Hilbert spaces, Fourier properties and analysis, randomized and low-rank approximations
- Asymptotics: quadratic expansions, central limit theorems, asymptotic normality, moment methods