COMS 4773 Spring 2024 Syllabus

1. Description

COMS 4773 ("Machine Learning Theory") is a graduate-level course on the theoretical study of algorithms for machine learning and high-dimensional data analysis. Topics include high-dimensional probability, theory of generalization and statistical learning, online learning and optimization, and spectral analysis.

2. Course information

- **Lecture times**: Tue/Thu 8:40am–9:55am
- **Lecture venue**: 451 Computer Science Building
- **Instructor**: Daniel Hsu

See course website https://djhsu.notion.site/COMS-4773-Spring-2024-ed665b71e9d4414b8de13db9c8d4d556 for up-to-date information and announcements.

3. Learning goals

The goal of this course (COMS 4773) is to equip you with tools to:

- Rigorously analyze machine learning problems and algorithms.
- Read and understand research papers in machine learning theory.

(Note: In this course, we won't really discuss application-oriented aspects of machine learning; you'll find more of that in COMS 4771.)

In his 1995 "Reflections After Refereeing Papers for NIPS", Leo Brieman offered four possible uses of machine learning theory, quoted below.

- **Comfort**: We knew it worked, but it's nice to have a proof.
- **Insight**: Aha! So that's why it works.
- **Innovation**: At last, a mathematically proven idea that applies to data.
- **Suggestion**: Something like this might work with data.

He also offers the following word of caution:

> Mathematical theory is not critical to the development of machine learning. *But scientific inquiry is.*

Theorems about machine learning algorithms and problems are useful for providing evidence of understanding and communicating ideas with clarity and precision. They are tools for engaging in scientific inquiry about machine learning.

4. Prerequisites

You should have mathematical maturity and be comfortable reading and writing mathematical proofs. We'll use a fair amount of probability and linear algebra; there'll also be a bit of convex analysis and algorithmic design and analysis (but not at a very advanced level).

You should have taken *some* proof-based course in mathematics, theoretical statistics, or theoretical computer science.

A course in machine learning (e.g., COMS 4771) is useful for understanding the motivation behind some problems and methods we discuss in this course. It is fine if you haven't taken such a course; just be prepared to do a bit of extra reading.

A previous course in learning theory (e.g., COMS 4252) is not required (and may, in fact, be somewhat redundant; see list of topics below).

If you have concerns about whether the course is suitable for you, please contact the instructor.

5. Topics

A tentative list of topics is as follows.

- Learning with expert advice
- Online convex optimization
- Multiarmed bandits
- Concentration of measure
- Statistical learning
- Uniform convergence
- Boosting and margins

- Approximation theory
- Fourier methods
- Gaussian concentration
- Spectral methods

There is some overlap with COMS 4252 and COMS 4774.

Related courses on machine learning theory at other institutions:

- Peter Bartlett's course at University of California, Berkeley
- Akshay Krishnamurthy's course at University of Massachusetts Amherst
- Matus Telgarsky's course at University of Illinois Urbana-Champaign

We won't directly follow any textbook, but the following may be useful references:

- Learning Theory from First Principles (Francis Bach)
- Foundations of Machine Learning (Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar)
- An Introduction to Computational Learning Theory (Michael J. Kearns and Umesh Vazirani)
- Mathematical Analysis of Machine Learning Algorithms (Tong Zhang)
- A Probabilistic Theory of Pattern Recognition (Luc Devroye, László Györfi, and Gábor Lugosi)
- Understanding Machine Learning: From Theory to Algorithms (Shai Shalev-Shwartz and Shai Ben-David)

6. Course requirements and assessment

You are expected to complete reading assignments, attend lectures and take careful notes. Lectures are unlikely to be recorded, so if you miss a lecture, please ask a classmate for notes.

- Homework assignments (60% of overall grade)
- Midterm exam (20% of overall grade), planned for week of March 4
- Reading project (20% of overall grade)

All assignments must be submitted as PDF documents compiled using TeX, LaTeX, or similar systems with bibliographic references (e.g., using BibTeX) included as necessary.

If you have not used LaTeX before, or if you only have a passing familiarity with it, it is highly recommended that you read and complete the lessons and exercises in The Bates LaTeX Manual.

Here is a LaTeX template you can use: template.tex.

7. Disability services

If you require accommodations or support services from Disability Services, please make necessary arrangements in accordance with their policies within the first two weeks of the semester.

8. Academic rules of conduct

You are expected to adhere to the Academic Honesty policy of the Computer Science Department, as well as the following course-specific policies.

8.1. Policies on plagiarism

Any work you submit must be written completely in your own words. Any material you quote or paraphrase must be clearly marked as such and properly attributed to its source.

8.2. Homework policy

(Homework policy is taken from http://www.cs.columbia.edu/~cs4252/mission.html.)

You are encouraged to discuss the course material and the homework problems with each other in small groups (2-4 people; 4 is an upper limit), but you must **list all discussion partners on your problem set**. Discussion of homework problems may include brainstorming and verbally discussing possible solution approaches, but **must not go as far as writing up solutions together; each person MUST WRITE UP HIS/HER SOLUTIONS INDEPENDENTLY. You may not collaborate with another student on writing up solutions or even look at another student's written solutions.** If your homework writeup resembles that of another student in a way which suggests that you have violated the above policy, you may be suspected of academic dishonesty.

You may consult certain outside materials, specifically lecture notes and videos of other classes, any textbooks, and research papers. You **may not consult any other materials, including solved homework problems for this or any other class. For all outside materials used, you must provide a detailed acknowledgement of the precise materials used. Whether or not you consult outside materials, you must always**

write up your solutions in your own words. If your homework writeup resembles any outside source in a way which suggests that you have violated the above policy, you may be suspected of academic dishonesty.

8.3. Exam policy

Exams must be completed individually. **Collaboration or discussion between students on exams is NOT PERMITTED. Outside references and sources CANNOT be used on exams.**

8.4. Consequences of policy violations

Violation of any portion of these policies will result in a penalty to be assessed at the instructor's discretion (e.g., a zero grade for the assignment in question, a failing letter grade for the course), even for a first offense.

9. Getting help

You are encouraged to use office hours and message board to discuss and ask questions about course material and reading assignments, and to ask for high-level clarification on and possible approaches to homework problems. If you need to ask a detailed question specific to your solution, please do so on the message board and mark the post as "private" so only the instructors can see it.

Questions, of course, are also welcome during lecture. If something is not clear to you during lecture, there is a chance it may also not be clear to other students. So please raise your hand to ask for clarification during lecture. Some questions may need to be handled "off-line"; we'll do our best to handle these questions in office hours or on message board.