EECS 598: Machine Learning Theory

Instructor: Wei Hu (vvh@umich.edu)

GSI: TBD

Time and location: Winter 2024

Description: When do machine learning algorithms work and why? How do we formally characterize what it means to learn from data? This course will study the theoretical foundations of machine learning. We will present the frameworks and rigorously analyze some of the most successful algorithms in machine learning that are extensively used. The course will prepare students for thinking rigorously about machine learning and doing research in a relevant area.

Prerequisites: This is an *advanced and theory-heavy* course. Students taking this course should have mathematical maturity and be comfortable with rigorous mathematical proofs. Familiarity with probability, multivariate calculus, and linear algebra is required. Knowledge of machine learning is strongly recommended but not strictly required.

Topics (tentative):

- Machine learning framework
- Concentration inequalities
- Generalization theory
    - Uniform convergence
    - Rademacher complexity
    - VC dimension
    - Margin theory
    - Covering numbers
- Optimization
    - Gradient descent
    - Optimization landscape
    - Convergence analysis
- Deep learning
    - Challenges in deep learning theory
    - Implicit regularization
    - Neural tangent kernel
- Online learning and bandits
    - Expert problem

- Online convex optimization

Grading:

- Homework assignments: 50-60%
- Final project (literature review and/or original research): 40%
- Class participation: 0-10%

Homework assignments: There are 4 homework assignments roughly equally spaced throughout the semester. The assignments consist of theoretical problems; there are no programming assignments.

Final project: Each student will complete a research-oriented final report, either individually or in a group of 2 people. There are two options:

1. Write a review of two papers on the same topic. The goal is to learn to read technical papers critically and present them in your own way. The review should include an in-depth exposition of the papers, summarizing the questions, results, significance, proof sketches, etc.
2. Write a review of one paper and do a mini follow-up project. The follow-up project can be any further investigation of the paper, such as generalization to another setting, simplification of the proofs, studying a different but related question, empirical study of the theory, etc.

A list of papers will be provided.

Textbook and references: There is no required textbook, but here are some useful sources and we will follow some materials from them in the course:

- Understanding Machine Learning: From Theory to Algorithms
- Foundations of Machine Learning
- Learning Theory from First Principles
- Lecture Notes for Machine Learning Theory