

CS6780 - Advanced Machine Learning

Fall 2023

Prof. Thorsten Joachims

Cornell University, Department of Computer Science & Department of Information Science

Time and Place

First lecture: August 22, 2023

Last meeting: November 30, 2023

Time: Tuesday/Thursday, 2:55pm - 4:10pm

Room: Gates G01 / Bloomberg 397

Exam: TBD

Project Report: TBD

Course Description

This course gives a graduate-level introduction to machine learning and in-depth coverage of new and advanced methods in machine learning, as well as their underlying theory. It emphasizes approaches with practical relevance and discusses a number of recent applications of machine learning in areas like information retrieval, recommender systems, data mining, computer vision, natural language processing and robotics. An open research project is a major part of the course.

In particular, the course will cover the following topics:

- **Supervised Batch Learning**: model, decision theoretic foundation, model selection, model assessment, empirical risk minimization
- **Decision Trees** : TDIDT, attribute selection, pruning and overfitting
- **Statistical Learning Theory** : generalization error bounds, VC dimension
- **Large-Margin Methods and Kernels**: linear Rules, margin, Perceptron, SVMs, duality, non-linear rules through kernels
- **Deep Learning**: multi-layer perceptrons, deep networks, stochastic gradient, computer vision
- **Large Language Models**: network architectures, fine tuning, preference learning from human feedback
- **Probabilistic Models**: generative vs. discriminative, maximum likelihood, Bayesian inference
- **Structured Output Prediction** : undirected graphical models, structural SVMs, conditional random fields
- **Latent Variable Models**: k-means clustering, mixture of Gaussians, expectation-maximization algorithm, matrix factorization, embeddings
- **Online Learning** : experts, bandits, online convex optimization
- **Causal Inference** : interventional vs. observational data, treatment effect estimation

The prerequisites for the class are: programming skills (at the level of CS 2110) and basic knowledge of linear algebra (at the level of MATH 2940) and probability theory (at the level of MATH 4710) and multivariable calculus (at the level of MATH 1920).

Enrollment is limited to PhD students.

Lectures

- 08/22: Introduction [slides] [slides 6-up]
 - What is learning?
 - What is machine learning used for?
 - Overview of course, course policies, and contact info.

Assignments and Exams

Homework assignments can be downloaded from [CMS](#) and they are submitted using the mechanism specified in the homework handout.

All assignments are due at the beginning of class on the due date. Assignments turned in late will be charged a 1 percentage point reduction of the cumulated final homework grade for each period of 24 hours for which the assignment is late. However, every student has a budget of 5 late days (i.e. 24 hour periods after the time the assignment was due) throughout the semester for which there is no late penalty. So, if you have perfect scores of 100 on all 4 homeworks and a total of 8 late days, your final homework score will be 97. No assignment will be accepted after the solution was made public, which is typically 3-5 days after the time it was due. You can submit late assignments via CMS.

Regrade requests can be submitted within 7 days after the grades have been made available using the mechanism specified in the homework handout.

Grading

This is a 4-credit course. Grades will be determined based on a written exam, a research project, homework assignments, and class participation.

- 40%: Exam
- 35%: Research Project
- 20%: Homework (~4 assignments)
- 5%: Class Participation (e.g., lecture, discussion)

To eliminate outlier grades for homeworks, the lowest grade is replaced by the second lowest grade when grades are cumulated at the end of the semester.

All assignment, exam, and final grades (including + and - of that grade) are roughly on the following scale: A=92-100; B=82-88; C=72-78; D=60-68; F= below 60.

For the research project, we will use peer review analogous to how scientific papers are reviewed for conferences and journals. This means that you will read and comment on other students work, which provides input for the TA's and the professor to determine the project grades. The quality of your reviewing also becomes a component of your own grade.

Students taking the class S/U do all work, except for the project, and need to receive at least a grade equivalent to a D to pass the course.

Academic Integrity

This course follows the [Cornell University Code of Academic Integrity](#). Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student in this course for academic credit will be the student's own work. Collaborations are allowed only if explicitly permitted. Violations of the rules (e.g. cheating, copying, non-approved collaborations) will not be tolerated. Respectful, constructive and inclusive conduct is expected of all class participants.