Spring 2023 **MATH 569 Statistical Learning**

Class time: MW 10:00am - 11:15am
Online via Blackboard Collaborate Ultra

Instructor
Maggie Cheng
E-mail: maggie.cheng@iit.edu
Tel: 312.567.8928
Office Hours: Wednesday 3:00pm - 4:30pm, online via Blackboard Collaborate Ultra

Teaching Assistant
Yuanxing Cheng
E-mail: ycheng46@hawk.iit.edu
Office hours: Monday 9:00am - 10:00am and Monday 11:30am - 1:30pm, RE 129

For online office hours:
Virtual RE129 (url: https://us.bbcollab.com/guest/704a8d4973ff4f15a00e1b6e62e9f5cc)
For instructions, use this link:
https://sites.google.com/iit.edu/virtual-math-learning-center/home#h.36h3gajmeken

Course Description
This course teaches modern statistical methods for learning from data, including regression, classification, kernel smoothing methods, and support vector machines; it also covers principles for model assessment, selection, and inference, as well as advanced topics such as learning in high-dimensional feature spaces.

Enrollment
Elective for AM MS, PhD plus MMF plus other majors.

Textbook(s)
Trevor Hastie, Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning, Springer (2009), 2nd edition, ISBN 978-0-387-84857-0.

Software: R, MATLAB, or Python

Prerequisites: MATH 350 Numerical Methods and MATH 474 or 475, or consent of the instructor

Objectives
1. Students will learn modern statistical techniques for modeling and drawing inferences from large data sets,
2. Students will learn to use visual and numerical diagnostics to assess the soundness of their models,
3. Students will become familiar with the computational requirements and compromises to be made in analyzing large data sets, and

4. Students will gain experience in analyzing real data sets and communicating their results.
5. Students will learn how to implement and use these numerical methods in MATLAB (or another similar software package),
6. Students will improve their problem solving skills in computational mathematics,
7. Students will improve their presentation and writing skills.

Assessment:

Homework (50%), midterm exam (50%), and an optional final exam.

Letter grades will be assigned based on grade-specific criteria:

Requirement for an A:

1. Maintaining a grade B throughout the semester.
2. Averaging 85% from midterm and final exams.
   Only the students who already have a B are eligible to participate in the final exam.

Requirement for a B: Getting an overall grade of 80% on homework and midterm, and the cumulative late time for all homework is no more than 48 hours.

Requirement for a C: Getting an overall grade of 65% on homework and midterm, and the cumulative late time for all homework is no more than 96 hours.

Requirement for a D: Getting an overall grade of 50% on homework and midterm, and the cumulative late time for all homework is no more than 120 hours.

Attendance: All students must attend at least 24 online lectures in order to maintain a grade of D or better.

Policy on late work:
Homework will not be accepted 24 hours after the deadline. Within 24 hours, it will be accepted and will count towards your total late time allowance.

Midterm: all students must take the midterm exam at the same time.

Offline mode: Students with full time employment can participate in an offline mode if they obtain a letter from their employer that verifies that the student is expected to be working during the lecture time.

Students in offline mode are not subject to the attendance requirement, but are still required to take the midterm exam at the same time with the rest of class.

Lecture Schedule: Two 75-minute lectures per week

Topics	Weeks
1. Overview of Statistical Learning	1.5
2. Linear Methods for Regression	2
3. Linear Methods for Classification	1.5
4. Basis Expansions and Regularization	2
5. Kernel Smoothing Methods	1.5
Midterm	
6. Model Assessment and Selection	2
7. Maximum Likelihood Inference and Bayesian Methods	2
8. Advanced Topics Selected from	2.5

- Boosting and Additive Modeling
- Neural Networks
- Support Vector Machines
- Nearest-Neighbor Methods
- Unsupervised Learning
- High-Dimensional Problems