

SUNY Albany
MAT592 Spring 2022
Machine Learning

Instructor: Felix Ye

Lectures: MoWe 3:00PM - 4:20PM in Massry Ctr for Business 141.

Instructor Office Hours: TBA.

E-Mail Address: xye2@albany.edu

Email will be a major line of communication between the student and the instructor. I will send urgent announcements and important information via email. Please check your university email regularly.

Web Page: Check the course page in blackboard regularly. Homework assignments, course announcements, and grades will be posted there.

Course Description: The primary goal is to provide students with the tools and principles needed to solve both the traditional and modern data science problems. In particular, the course covers a wide variety of topics in machine learning. It introduces the key terms, concepts and methods in machine learning, with an emphasis on developing critical analytical skills through hands-on exercises of data analysis tasks. In addition, the students will practice basic programming skills to use software tools in machine learning. The programming language in this course will be Jupyter notebook.

All latest jupyter notebook can be downloaded in <https://github.com/yexf308/MAT592.git>. I will keep updating this git repository as class progresses, so please fetch the update regularly. I will also leave a copy (not most updated) version in blackboard.

Prerequisite: vector calculus, probability, linear Algebra, optimization and some experience on programming.

Textbook: Probabilistic Machine Learning: An Introduction

You can download the draft of the latest version: <https://probml.github.io/pml-book/book1.html>. I will loosely follow this book, however, I will put some additional material into it.

Grading Policy:

Homework	70%
Take-home Final	30%

Incomplete grade: This class will not give any incomplete grade. If the work cannot complete in the current semester, the student can choose to retake this class in the following year.

Exams: There will be one take-home final. The detail will be given at the end of the semester.

Homework: Homework assignments will be assigned each month. There are 3 sets of homework in total.

Late assignment turn-in is not permitted. Any assignment turned in after the deadline will NOT be graded.

Attendance: Although attendance will not be taken, I strongly encourage you attend and participate in every lecture. This is one of the best ways to ensure success in the course.

Academic Misconduct: The strength of the university depends on academic and personal integrity. In this course, you must be honest and truthful. Ethical violations include cheating on exams, plagiarism, reuse of assignments, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition.

In addition, specific ethics guidelines for this course are as follows: Students may discuss homework. However, all solutions MUST be written up and submitted individually. The same rules apply to computer programs. Basic ideas may be discussed but detailed codes should not be copied or shared. Finally, exams must represent the result of individual effort and communication is permitted only with the instructor.

Report any violations you witness to the instructor. You may consult the associate dean of student affairs and/or the chairman of the Ethics Board beforehand.

Tentative Course Outline and Schedule:

- Week 1 (Jan 19): Machine Learning: Overview.
- Week 2 (Jan 24 & Jan 26): Review of Pre-requisite: linear algebra, optimization and probability.
- Week 3 (Jan 31 & Feb 2): Linear regression. Homework 1 assigned.
- Week 4 (Feb 7 & Feb 9): Logistic regression.
- Week 5 (Feb 14 & Feb 16): Support Vector Machines. Homework 1 due.
- Week 6 (Feb 21 & Feb 23): More on Support Vector Machines. Kernel method. Homework 2 assigned.

- Week 7 (Feb 28 & Mar 2): KNN classification.
- Week 8 (Mar 7 & Mar 9): Naive bayes classification. Homework 2 due.
- Spring Breaks.
- Week 9 (Mar 21 & Mar 23): K-means clustering. Homework 3 assigned.
- Week 10 (Mar 28 & Mar 30): Gaussian mixture model.
- Week 11 (Apr 4 & Apr 6): Principal component analysis.
- Week 12 (Apr 11 & Apr 13): Decision tree. Homework 3 due. Take-home final assigned.
- Week 13 (Apr 18 & Apr 20): Deep Neural Network.
- Week 14 (Apr 25 & Apr 27): Deep Neural Network.
- Week 15 (May 2): Summary. Take-home final due.