

DATASCI 347: Machine Learning

Spring 2026

General Information

Class Meeting Times: Mondays and Wednesdays, 11:30 AM - 12:45 PM, Psychology Building 225

Instructor: Ruoxuan Xiong, Psychology and Interdisciplinary Sciences Building 581, ruoxuan.xiong@emory.edu

Office Hours: Wednesdays 3:00 PM - 4:00 PM, Psychology and Interdisciplinary Sciences Building 581

Course Description

This course introduces students to the field of machine learning, a critical toolset for analyzing and interpreting complex data sets across diverse domains, including biology, finance, marketing, and astrophysics. Students will explore foundational modeling and prediction techniques widely used in machine learning, artificial intelligence, and data science, with a focus on both practical applications and the statistical principles underlying these methods.

Topics covered include:

1. Supervised Learning (Regression and Classification): Linear regression, logistic regression, K-nearest neighbors, linear and quadratic discriminant analysis, regularization methods (Ridge, LASSO, and Elastic Net), and tree-based methods.
2. Model Evaluation and Resampling: Bias-variance tradeoffs, cross-validation, and bootstrapping.
3. Unsupervised Learning: Principal components analysis, clustering, and k-means algorithms.
4. Reinforcement Learning: Markov decision process, value iteration, policy iteration, model-based RL, and policy gradient.
5. Advanced topics: Neural networks, transformers, diffusion models, and foundation models

By the end of this course, students will gain both theoretical knowledge and practical skills to apply machine learning techniques to real-world problems.

Grading

You are responsible for keeping up with all announcements made in class and for all changes in the schedule that are posted on the class website.

The grade will be based on the following:

- Homework 30%
- Exam 30%
- Final project 35%
- Participation 5%

Homework

There will be a total of 3 homework assignments. The homework assignments consist of both theoretical and empirical questions. The main statistical software used in class is Python.

The homework assignments are done in groups. The group size is up to four students. You should choose the same group for all homework assignments. The homework will be graded on a 100-point scale. You have a total of three free late days for all homework assignments as a group. You can use at most two late days for one homework assignment.

Homework is assigned and due as follows:

- Homework 1 handed out on Jan 21, 2026 , due on Feb 11, 2026
- Homework 2 handed out on Feb 11, 2026 , due on Mar 4, 2026
- Homework 3 handed out on Mar 4, 2026 , due on Apr 1, 2026

Exam

The exam will be a take-home exam. The exam will be handed out at Apr 8, 2026 00:00 am and due at Apr 12, 2026 11:59 pm. You can choose any 24 hours in between to complete. There is no make-up exam.

Course Project

The goal of the course project is to prepare you for some project experience in machine learning. By the end of the project, we hope that you will have gained some hands-on experience in applying ML to a real-world problem, or learned some research frontiers in machine learning. We will provide a sample list of datasets and papers. You have two options to complete the project. The first option is to pick a dataset that interests you, and apply the knowledge we have gained this semester to analyze this dataset. The second option is to replicate a research paper and explore the possible extensions/improvements of the paper.

The course project is done in the same group as the homework.

There is a project proposal presentation on **Mar 18, 2026**. Each group needs to prepare a five-minute presentation that includes all the group members (up to four students) and the topic of your group.

There are final project presentations on **Apr 22, 2026** and **Apr 27, 2026**. Each group needs to prepare a ten-minute presentation that includes the motivation, setup, and results of the project. Before the full project presentation, we ask that you set up a publicly available GitHub repository about your work, along with detailed documentation about how to use the code repository and what findings you currently have about the project.

We expect when each group presents, other groups will provide critical feedback, which will be counted toward the participation in this course.

Finally, by **May 6, 2026**, refine the GitHub repository and the accompanying documentation.

Prerequisites

QTM 220 Regression Analysis or equivalent.

Textbook

You will only be tested on the material presented in lectures, and learned through the homework. Here are some useful references for the course material:

- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*. Vol. 112. New York: springer, 2013.
- Hastie, Trevor, Robert Tibshirani, Jerome H. Friedman, and Jerome H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. New York: springer, 2009.

Honor Code

All students enrolled at Emory are expected to abide by the Emory College Honor Code. Any type of academic misconduct is not allowed, which includes 1) receiving or giving information about the content or conduct of an examination knowing that the release of such information is not allowed and 2) plagiarizing, whether intentionally or unintentionally, in any assignment. For the activities that are considered to be academically dishonest, refer to the Honor Code:

<http://catalog.college.emory.edu/academic/policies-regulations/honor-code.html>.

Accessibility and Accommodations

As the instructor of this course, I endeavor to provide an inclusive learning environment. I want every student to succeed. The Department of Accessibility Services (DAS) works with students who have disabilities to provide reasonable accommodations. It is your responsibility to request accommodations. In order to receive consideration for reasonable accommodations, you must register with the DAS at <http://accessibility.emory.edu/students>. Accommodations cannot be retroactively applied, so you need to contact DAS as early as possible and please contact me as early as possible in the semester to discuss the plan for implementation of your accommodations.

For additional information about accessibility and accommodations, please contact the Department of Accessibility Services at (404) 727-9877 or accessibility@emory.edu.