Mathematical Foundations of Machine Learning

ECE/ISYE/CS/CSE 7750

Fall 2026 Info

Instructor

Sara Fridovich-Keil
Email: sfk@gatech.edu
Office hours: Tuesdays 4-5 pm (Coda C1003 Adair, starting September 8), Thursdays 4-5 pm (Coda C1003 Adair, starting September 3)

Please use my office hours to ask questions about course material and beyond (e.g., research); questions specific to homework problems are best asked at TA office hours (see below).

The best online way to contact me with questions about the course is through Ed (see below); please use email as sparingly as possible.

Teaching Assistants

TAs will also hold weekly office hours, which are the best place to ask questions about homework problems.

Homework

Homework will be submitted online through Canvas/Gradescope (exact instructions will provided when the time comes). You have **5 late days in total** that you can use throughout the semester. Beyond this, late homework will get zero credit. Over the course of the semester, **the maximum number of homework points that you can earn is $(N-1) \cdot 100$, where N is the number of assignments** – this serves a similar role to allowing you to drop one homework assignment, but should encourage you to still submit a partially completed one (and avoid panic if there is occasionally a problem that you do not finish in time.)

There will be some basic programming on the homeworks. These will be completed in Python using Jupyter notebooks. If you have not programmed in Python in the past, this is your chance to pick up some basic skills. The programming problems are basic simulations and evaluations on datasets; they are straightforward enough that I have confidence that anyone with any type of programming experience will be able to complete them. Make sure you have a working Jupyter notebook installation as well as capability to install additional open-source Python packages as needed (Homework 0 will point you to some resources for this).

The homework assignments will be hard; many of them will require significant amounts of time and effort to complete. But this is really where most of the learning takes place. You will get out of the assignments what you put into them. Students who complete all of the assignments in full will be rewarded with a deep understanding of the role that linear algebra and statistics play in machine learning. Effectively, homework is worth much more than 50% of your grade. In teaching many courses over the years, the instructors **have never seen a case where a student does not put effort into the homework assignments but does well on the exams**.

Students are *strongly* encouraged to discuss homework problems with one another. Some of the office hours will be specially designated as "homework discussion hours" where the first half of the office hour will be attended by students only; the purpose of this is to encourage you to help each other out without instructor involvement. However, **each student must write up and turn in their own solutions written in their own words**. Cases where solutions appear to be identical or nearly identical will be immediately referred to the Office of Student Integrity.

Unauthorized use of any previous semester course materials, such as tests, quizzes, and homework, is prohibited in this course. Furthermore, redistributing materials from this semester is also prohibited. For any questions involving these or any other Academic Honor Code issues, please consult me or www.honor.gatech.edu.

Homework grading: Homework for this class will be self-graded using provided rubrics. Self-grading deadlines will be roughly a week after the homework assignment is due, and will also be submitted via Canvas/Gradescope. The TA will pick *1-2* HW problems at random for each HW, give you feedback (e.g., on whether you are under or over-reporting your grade), and adjust the self-grade on those problems if necessary.

Lecture

Lectures are Tuesday and Thursday from 5-6:15 pm in Scheller 300. Lectures will *not* be recorded, so **I strongly recommend that you attend lecture in person to keep up with the class.** Extensive course notes are provided and will be uploaded before each lecture. These will help you catch up on an occasional missed lecture; if you have further questions about a missed lecture, please see me at office hours.

Exams

Exams will be in-person during specified time frames. If you expect to miss an exam, please contact me as soon as you realize this so we can make alternative arrangements.

Online course materials

Lecture schedule, notes for the lectures, and homework schedule will be posted on Canvas. Students can use Canvas to access supplementary materials for the lectures (e.g., demos), homework assignments, and review their grades.

We plan to make exclusive use of Ed (accessible via Canvas) to make announcements and answer questions: https://edstem.org/us/courses/101047/discussion . Ed is a great platform for you to work with your fellow students to discuss problems, form study/project groups, etc. **Please direct any questions you might have to Ed.** Unless your question is personal in nature, please do not make a private post — if you have a question you are probably not the only one, and other students may benefit from seeing the discussion. Students can be anonymous to each other (if they wish) but will not be anonymous to instructors.

I also encourage students to answer questions. The enrollment in this class is fairly large, so there will be a great economy of scale here. If you *answer* at least 10 student questions *substantively* by the end of the semester, there is an opportunity for an extra point of credit. Please keep this in mind as you engage with Ed throughout the semester.

Text

There is no required text. Extensive course notes will be provided. From time to time, I will point you to books, videos, and notes by others that are good resources for the material we are talking about.

Outline

The outline below should be treated as an approximation; it is subject to (small) changes.

1. Linear Representations
   (a) linear vector spaces, linear independence
   (b) norms and inner products
   (c) linear approximation
   (d) basis expansions
2. Regression using Least Squares
   (a) regression as a linear inverse problem
   (b) the least-squares problem
   (c) ridge regression
   (d) regression in a Hilbert space, representer theorem
   (e) reproducing kernels, kernel regression, Mercer's theorem

3. Solving and analyzing least-squares problems
    (a) the Singular Value Decomposition (SVD) and the pseudoinverse
    (b) stable inversion and regularization
    (c) gradient descent, stochastic gradient descent, and conjugate gradients
    (d) matrix factorization
4. Estimation and Prediction
    (a) review of probability
    (b) Gaussian estimation and Gaussian random processes
    (c) parametric inference: the MLE and the MAP
    (d) classification and logistic regression
    (e) empirical risk minimization