7750: Mathematical Foundations of Machine Learning

A proof-based, graduate-level grounding in linear algebra and probability

Course syllabus

Course summary and philosophy

This course for beginning graduate students develops the mathematical foundations of machine learning, rigorously introducing students to modeling and representation, statistical inference, and optimization. The class will rigorously build up the two pillars of modern machine learning: linear algebra and probability. The class is proof-based, and much of our development will be theoretical, although there will also be exposure to simulating algorithms to see when they work (or don't work). The eventual focus is not necessarily to understand cutting-edge machine learning (although we will touch upon this), but the mathematical principles on which these ideas are built.

Learning outcomes

Upon successful completion of the course, you will have learned:

(a) The linear algebraic principles behind modeling function/hypothesis classes, with exposure to both finite and infinite dimensional modeling techniques.

(b) The probabilistic principles based on which we can perform statistical estimation with our models given data.

(c) Some basic principles that govern the design and analysis of optimization algorithms used to fit models to data.

The most important takeaway for some of you might be to recognize that these ideas can help in designing new, principled machine learning methodology, or conversely, to recognize the immense opportunity that exists to place several modern machine learning techniques on a rigorous footing.

Preparation and prerequisites

Students will be expected to have a working knowledge of probability and statistics, linear algebra and multivariable calculus (at the level of MATH 1553 and 2551 or equivalent), basic optimization, and proficiency with Python programming. The most important prerequisite is that you are open to learning (if you don't already) how to write rigorous proofs. Having taken a rigorous, proof-based undergraduate course will prove very helpful.

We will try our best to bring you up to speed with some of the prerequisites by using some auxiliary handouts, reviews during lecture, and optional homework (HW0). Given the breadth of machine learning and the fact that we will try to cover topics with rigor, **the class will be proceed at a fast pace**. Students should expect to do extra work in proportion to the amount of background that they are missing.

Course staff

Instructor:

- **Ashwin Pananjady (ashwinpm@gatech.edu):** Ashwin is a faculty member with a joint appointment between ISyE and ECE at Georgia Tech. He got his Ph.D. in EECS at UC Berkeley, and then spent a semester as a postdoctoral research fellow at the Simons Institute for the Theory of Computing. He studies high-dimensional statistical methods as well as optimization theory and algorithms, and works on problems arising from the intersection of these disciplines with machine learning, reinforcement learning, and signal processing.

TAs: TBA. Will be assigned within the first week of classes.

Scheduled live meetings

Lecture (Weber Space Science and Technology Lecture Hall 1):	3.30-4.45pm MW
Instructor OH:	5-5.30pm M and 8-9am Th (online)
Problem solving session and TA OH (Location TBD):	5.30-7pm F

Special announcements: There are riders to the regular schedule above:

- There will be no lecture on the day of a midterm.
- No lecture on Sep 5 due to Labor day.
- There will be no lectures on Oct 17 and 19, and no OH on Oct 17, owing to Fall break. Have fun and stay safe!
- There will be no OH or lecture during the Thankgiving week of Nov 21, only a midterm on Nov 21.
- No meetings or OH during dead week (Dec 5-9).

Grading

The course will be graded on the following components:

- **Homework** (50%): Homework will be roughly fortnightly, and there will be 6-7 homework assignments in total. Further details on homework are listed below.
- **Midterms** (50%): There will be two midterm exams. Tentatively, the first will be conducted during the week of Sep 26. The second will be conducted on November 21 (Monday before Thanksgiving weekend). There is no final exam for this class.

The students' final grade will be assigned as a letter grade according to the scale:
A: $90 - 100\%$ B: $80 - 89\%$ C: $70 - 79\%$ D: $60 - 69\%$ F: $\leq 59\%$

The instructor may exercise the option to "curve" midterm scores if the midterm were determined to be more difficult than intended. If exercised, this would only result in grades being adjusted *higher* for all students.

Course expectations and guidelines

Academic integrity Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code, please visit `http://www.catalog.gatech.edu/policies/honor-code`. Any student suspected of cheating or plagiarizing on a quiz, exam, or assignment will be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations. This is a graduate class; we expect that you are here to learn something new and that you will be conscientious and conduct yourself with courtesy and integrity.

Any referencing of resources from past offerings of this class (that are not provided as reading material in this offering) is strictly prohibited. Redistributing materials for this course and/or using external sites for assistance (e.g. contributing to test banks, CourseHero, Chegg, or similar sites) is prohibited.

Collaboration and group work Students are strongly encouraged to discuss homework problems with one another. However, *each student must write up and turn in their own solutions, written in their own words/consisting of their own code. Cases where written solutions or code appear to be identical or nearly identical will be immediately referred to the Office of Student Integrity.*

Absences/late submissions Out of fairness to the entire class and to ensure that we keep to our HW schedule, *late submission of homework will not be accepted* (but note that your lowest HW score will be dropped, see below). Absence at midterm exams will not be accepted in the absence of a prior agreement between the student and instructor. In particular, *excused absences for midterms* include illnesses, religious observations, career fairs and job interviews. A student who expects to miss a midterm due to an excused absence should contact the instructor as soon as possible so that the instructor can make alternate arrangements. Such arrangements could be taking the midterm at an alternate time or adjust the grading allocation depending on the circumstances.

Etiquette for questions: Aside from questions you have in our in-person meetings, please put all technical questions on Piazza. We will make an effort to curate this material with separate threads for each problem so that we enable peer-to-peer question answering to the best extent possible. The instructor and TAs will answer questions on Piazza. We will try our best to answer your question during the day in which it is asked, but please help each other out on this front. **If you are able to provide informative answers to peer questions on Piazza, then we will assign bonus credit.**

For logistical questions, please create a separate thread on Piazza between you and the TAs (and if necessary, the instructor) and ask it there. If your question is urgent and not answered on the private Piazza post, please email the course staff with subject line beginning with [7750].
Note: The instructors will be most responsive on Piazza, not email. If you have an urgent question that cannot be answered by the class, please create a private post on Piazza first; only send email if absolutely necessary.

Accommodations for students with disabilities If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also email the instructor ASAP to discuss your learning needs.

Student-faculty expectations agreement At Georgia Tech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class. See `http://www.catalog.gatech.edu/rules/22` for an articulation of some basic expectation that we can have of each other.

Support for student health and well-being: This is still an unprecedented time, so please be kind to yourself. We are all under a lot of stress after spending over two years in a pandemic, and I applaud your resolve and drive in choosing to learn something new this semester. Make sure to eat well, exercise, and reach out to your support system or the course staff if you need to.

Homework

Homework will be assigned approximately biweekly, with deadlines typically on Tuesday 11.59pm ET. *Homework must be turned in via Canvas. Late submissions will result in zero credit.* Each homework assignment has a maximum of 100 points, and the best $N-1$ out of the N homeworks will be evaluated. Thus, *the maximum number of homework points you can earn is equal to* $(N-1)\cdot 100$. *This also means that you can drop one out of the N homeworks if you like, or alternatively, give yourself some breathing room to submit a partially completed one.*

In keeping with the course philosophy, each homework assignment will be largely mathematical/conceptual ($\sim 80\%$). Homework problems will build on lecture material in substantive ways, and complement your understanding of the concepts. The remaining $\sim 20\%$ of the homework will use Jupyter notebooks for basic simulation and evaluations on datasets. Even these will be of the conceptual bent. *Students will need coding familiarity with Python to do the homework. Students will need a working Jupyter notebook installation as well as capability to install additional open-source Python packages as needed.*

HW submission guidelines. For each assignment, you will need to submit:

- A pdf named: 'Lastname_Firstname_HWx_report', with x replace by the HW number. In this report file, you need to summarize ALL your written answers to questions that did not involve coding/simulation. If you choose to hand-write the report, please emphasize legibility.
- If the HW includes a coding component, please also submit a zip file named 'Lastname_Firstname_HWx.zip': In this zip file, there should be a folder 'Lastname_Firstname_HWx', which should include all your code files. We strongly recommend submitting code using well-documented Jupyter notebooks; this will make it easier both for you and us.

Also ensure that:

(a) Your code doesn't have directory dependencies on your computer. Your code should be executable without path issues.

(b) Any data read by your code is included in your submitted zip file, so that the Jupyter notebook is runnable.

Solutions and grading. Solutions for the HW will be released immediately after HW deadline for the distance learning sections. Given that this is a large graduate class, our primary mode of grading will be a **self-grade** of your HW. In particular, you will be expected to cross-check your answers with the released solutions, and hand in on Canvas (one week after the solutions are released) a detailed self-grade for the HW. You can also indicate any lingering questions you have on this form. Your HW and self-grade will be checked by the course staff to ensure that credit is being assigned appropriately, and we will contact you to adjust if we see any discrepancies. The last HW for the semester will be graded by the course staff, not self-graded.

Rough schedule for HW: The tentative schedule for HW is below. All deadlines will be 11.59pm ET, with submission on Canvas.

- HW0: Due on Friday Aug 26, will not be graded. Only for you to assess your background and review some undergraduate material. Bonus credit 2 points. Preparing using this HW is highly recommended, it will pay off throughout the semester.
- HW1: Released Wed Aug 24, due on Tue Sep 6.
- HW2: Released Wed Sep 7, due on Tue Sep 20.
- HW3: Released Wed Sep 21, due on Tue Oct 4.
- HW4: Released Wed Oct 5, due on Fri Oct 14. This will be a shorter HW in view of Fall break.
- HW5: Released Wed Oct 19, due Tue Nov 1.
- HW6: Released Wed Nov 2, due Tue Nov 15.
- HW7: Released Wed Nov 22, due Fri Dec 2.

Note: The last HW for the class is the final deliverable, and is due on a Friday instead of a Tuesday.

Course materials

Course materials will be made available on Canvas and Piazza. Homework assignments and solutions will also be posted here as they are made available.

The instructor and TAs will make exclusive use of Piazza to make announcements and answer questions. Piazza is a great forum to discuss problems, find study groups, etc. Please direct any questions you might have about the course to Piazza. Unless your questions are personal in nature, please do not make private posts: if you have a technical question you are probably not the only one, and other students may benefit from seeing the discussion!

Resources: The course will cover several deep fields of study and each by itself could make up an entire course. This course is designed as a introductory gateway to these areas. As a result, we will not be using a single "required" textbook, but drawing from multiple references. The main reference for the class will be excellent course notes developed by other lecturers in previous offerings of the class, which will be edited and released as the class progresses.

References to some useful material are below (those available freely online—usually on the authors' websites—are marked F). Some lectures will draw on material from these references and we will point it out as and when we do so.

Linear algebra and least squares

- Numerical Linear Algebra by Trefethen and Bau.
- Introduction to applied linear algebra by Boyd and Vandenberghe (F)
- Essence of linear algebra playlist by 3Blue1Brown

Probability and statistics

- Introduction to Probability by Bertsekas and Tsitsiklis
- Elementary Probability for Applications by Durrett
- All of Statistics by Wasserman

Matrix calculus

- The matrix cookbook (F)
- These notes on matrix calculus

Statistical learning

- An introduction to statistical learning by James, Witten, Hastie, Tibshirani (F)
- Foundations of Machine Learning by Mohri, Rostamizadeh, Talwalkar (F)

Optimization

- Convex optimization by Boyd and Vandenberghe (F)
- Lectures on Modern Convex Optimization by Ben-Tal and Nemirovski
- First-order and Stochastic Optimization Methods for Machine Learning by Lan

Outline of class

The class will roughly be split into three modules: Representation, Optimization, and Estimation/Prediction. We will attempt to cover the following (approximately ordered) list of topics:

Part I: Representation

- Vector spaces
- Linear approximation and basis expansions
- Regression using least squares
- Formulating linear inverse problems
- Regression in Hilbert spaces
- Reproducing kernels and Mercer's theorem

Part II: Optimization

- Solving least squares problems
- Stable inversion and regularization
- Singular value decomposition
- (Stochastic) gradient descent
- Conjugate gradient
- Matrix factorization

Part III: Estimation and prediction

- Basic statistical principles: Generative modeling, Bayesian and frequentist thinking, maximum likelihood and MAP.
- Bias-variance trade-off, and validation on holdout data
- Predicting probabilities: classification and logistic regression
- Empirical risk minimization