Summary of course

At the heart of most machine learning and AI applications today – like advertisement placement, movie recommendation, and refining models from human feedback – is an optimization engine trying to provide the best decision with the information observed thus far in time, i.e. the problem of *online learning*. To solve these problems, one must make online, real-time decisions and continuously improve the performance with the arrival of data and feedback from previous decisions. The course aims to provide a foundation for the development of such online methods and for their analysis. We will discuss fundamental principles for learning from an unknown environment, limited feedback, and learning with dynamic, long-term consequences.

Learning outcomes

Upon successful completion of the course, students will be able to:

- Understand where online learning is applicable in many real-world scenarios.
- Develop algorithms that combine partial information as best as possible to make online decisions.
- Understand how exploration of decision space and exploitation from historic data must be prioritized to be able to reach optimal decisions.
- Understand the *dynamic-programming* principle of sequential decision-making when decisions have long-term consequences, and appreciate principles for learning in such long-term environments.

Prerequisites

The listed prerequisites are ISYE 3133 (Optimization) and ISYE 2027 (Probability). This is an advanced course, and there is a lot of subtle mathematics that we will cover to understand the basics of online machine learning. We will expect some mathematical maturity in linear algebra, set theory, and analysis as well. Assignment 0 covers the background needed for this class including basics of probability, set theory, analysis and optimization. If a student gets below 50% in Assignment 0 or does not attempt Assignment 0 meaningfully, they are recommended to drop the class. Further, assignments in this course are mathematical as well as programming-based and require a working knowledge of Jupyter notebooks and Python.

Course staff

Instructor - Vidya Muthukumar (vmuthukumar8@gatech.edu): Vidya is a faculty member with a joint appointment between ECE and ISyE at Georgia Tech. She received

her 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 for the program on Theory of Reinforcement Learning. Her research interests are in game theory, online learning and statistical machine learning. She current works on designing learning algorithms that provably adapt in strategic environments and uncovering fundamental properties of modern ML models.

UTA/MTA: TBD

Scheduled live meetings

Lecture (TSRB 132): 9:30-10:45 am Tuesdays and Thursdays

Instructor office hour (11 am-12 pm Tuesdays): George Tower 1114

Homework/exam discussion office hour: venue and weekly time TBD

Grading

The course will be graded on the following components:

- **Assignment 0 (2% bonus)**: Assignment 0, which will be released on **Aug 25** and due **Sep 1**, will be significantly shorter than the other assignments and will cover prerequisite material for the class. If a student gets below 50% on Assignment 0 or cannot meaningfully attempt Assignment 0, they are recommended to drop the class.
- **Homeworks (30%):** Homework will be roughly fortnightly (with breaks or longer timelines during midterm weeks), and there will be 4 homework assignments in total. Further details on homework are listed below.
- **Two midterms (40% in total):** There will be two midterm exams (each of which contribute to 20% of your grade) that will test conceptual and technical knowledge of the foundational components of online learning and decision-making. The first midterm will be held on **Thursday, Sep 24** and the second midterm will be held on **Tuesday, Oct 27** --- both during the regular lecture hour.
- **Final exam (30%):** The final exam will cover the entire course, and will be held on **Monday Dec 14** in accordance with Georgia Tech's guidance for exam scheduling.

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: <= 59%

The instructor may exercise the option to ``curve" midterm or final exam scores if determined that midterm/final exam was more difficult than intended. This would result in grades being adjusted *higher* for all students, if at all.

Homework

Homework will be assigned approximately biweekly. Some general guidelines for homework policy are provided below.

- **Submission guidelines:** The assignments have to be submitted online via Canvas – either with typed up solutions (word/LaTeX) or images of handwritten notes. Points will be taken off if the handwritten notes are not legible.
- **Coding-based questions:** For coding based questions, you have to submit the output as well as the code. Students will need a working knowledge of Python (basic packages) and Jupyter notebooks to solve the coding-based questions.
- **Collaboration policy:** You are encouraged to discuss homework/assignments problems with your fellow students. However, your final answers should be based on your own understanding unless it is a group assignment (which will be announced on Canvas if it is the case). Copying others' work is *not* acceptable and violates the honor code. At the top of your submission, please write the names of students you have discussed the assignments with if any. In case answers are found copied from another student, both the students (the one who allowed copying, and the one who copied) will both get a 0 on the entire assignment. Repeat offenses of this type risk a D/F grade for the entire course. Our collaboration policy is in accordance with Georgia Tech's Academic Honor Code: see more details under the section **Course expectations and guidelines** below.
- **AI use policy (Limited Generative AI use permitted):** Use of Generative AI (e.g. ChatGPT, Gemini, Claude, Microsoft Copilot, etc.) for homework is permitted in this course in a *very limited sense*, i.e. within instructor-approved boundaries. You can use AI to assist with English writing in your solutions or formatting of your solutions. You *cannot* use AI to generate math solutions or write code in any capacity (including checking your solutions/code). *Any usage of AI must be transparent and documented in a required AI usage statement with each submission. Failure to follow these guidelines – including using AI when it is not permitted or failing to disclose its use – may be considered a violation of Georgia Tech's academic integrity policies.* The statement must include:
    - Tool used and date of access

    - The input (prompt) that you provided
    - A copy of the output
    - A description of how you used or edited the AI-generated content
- **Regrade requests:** will be accepted via private Piazza post or Gradescope to the instructor and the TA within 5 days of the graded assignment/exam being returned. Requests must include (1) a PDF of the original submission and (2) a LaTeX-produced PDF detailing which problems were graded incorrectly and an argument that the submitted solution is indeed correct. Regrades may only be requested if it is believed that a correct answer was marked as incorrect, not because insufficient partial credit was given to an incorrect or partially correct solution. Keep in mind that if you request a regrade for a part of the relevant assignment/exam question, there is a good chance that the entire question will be regraded, not just the part that you believe to be graded incorrectly.
- **Late submissions:** For every 12 hours that the assignment is submitted late, the grade will be reduced by 10% (linearly interpolated), without exceptions. The only official extensions for late assignments will be granted with a letter from the Dean of Student's Office to be received before the assignment is due. See more details under the section **Course expectations and guidelines** below.

Course materials

A course webpage will be made publicly available. This page will provide general course information, links to typeset lecture notes, and links to homework assignments. Homework assignments and solutions will also be posted in canvas 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 problem 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 question you are probably not the only one, and other students may benefit from seeing the discussion!

Textbook/class resources

Typeset notes accompanying each lecture will be posted after the lecture has concluded. Supplemental notes that were originally developed by Dr. Swati Gupta and Jad Salem will also be posted; these are optional reading.
This is a fairly new and modern area of machine learning and optimization; the following are excellent references.

- Nicolo Cesa-Bianchi and Lugosi: Prediction, Learning and GamesLinks to an external site.
- Elad Hazan: Introduction to online convex optimizationLinks to an external site.
- Tor Lattimore and Csaba Szepesvari: Bandit algorithmsLinks to an external site.
- Dimitri Bertsekas: Dynamic Programming and Optimal ControlLinks to an external site. (volumes 1 and 2)
- Martin Puterman: Markov Decision Processes: Discrete Stochastic Dynamic ProgrammingLinks to an external site.

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 this link. 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.

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. Cases where solutions 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, *late submission of homework, or absence at the midterm exam, will not be accepted in the absence of a prior agreement between the student and instructor*. In particular, *excused absences* include illnesses, religious observations, career fairs and job interviews. In the event than an excused absence such as above prevents a student from submitting an assignment, their homework grade will be calculated on a prorated basis. A student who expects to miss the 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.

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 or disabilityservices.gatech.edu as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter.

Please also email the instructor as soon as possible in order to set up a time 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 https://catalog.gatech.edu/rules/22/ for an articulation of some basic expectation that we can have of each other.

Outline of course

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

Part I: Online "full-information" learning

- Principles for prediction from time-varying/adversarial data and learning with expert advice
- The multiplicative weights algorithm and applications in solving linear programming, zero-sum games

Part II: Multi-armed bandits

- The exploration-vs-exploitation principle in stochastic bandits
- Foundational bandit algorithms: UCB and Thompson sampling
- Applications to recommendation systems and the basic model for Reinforcement Learning from Human Feedback (RLHF)

Part III: Special topics in reinforcement learning (RL)

- Dynamic programming and optimal control
- Basic principles of reinforcement learning and applications to operations management
- (if time) Discussion of how RL is used in large language models, e.g. verifiable rewards (RLVR)