CS 4641 Syllabus

CS 4641: Machine Learning; Section D; 3 Credits

Fall 2026

Instructor Information

Instructor: Jesse Thomason

Email: thomason@gatech.edu

TA: Davis Clark

Email: dclark304@gatech.edu

General Course Information

Description

This course introduces techniques in machine learning with an emphasis on algorithms and their applications to real-world data. We will focus on machine learning methods, which are organized into three main course objectives.

1. *Basic math for data science and machine learning*
2. *Unsupervised machine learning for data exploration*
3. *Supervised learning for predictive data analysis*

Prerequisites for this course include (1) basic knowledge of probability, statistics, and linear algebra; (2) basic programming experience in Python.

Course Learning Outcomes

- Introduce to you the pipeline of Machine Learning
- Help you understand major machine learning algorithms
- Help you learn to apply tools for real data analysis problems
- Encourage you to do research in data science and machine learning

In addition to the technical content, this class includes the following learning objectives:

- Structuring a task into a machine learning work flow

- Collaborating effectively on team projects
- Conducting peer evaluation in a constructive format
- Communicating technical content in a concise and effective manner

Required Course Materials

No textbook will be required for this course. You may find the following books helpful:

- Pattern Recognition and Machine Learning, by Christopher Bishop
    - Your instructor still has this one on his bookshelf at home and has consulted it regularly since he got it in undergraduate ML circa 2011.
- Deep Learning, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville
    - This one is made available in full online by the authors.

Course Schedule

Lecture topics are subject to change as the semester goes on.

W	Date	Topic	HW	Quiz	Project
1	Aug 25	Course Overview			
	Aug 27	ML Notation and Terminology			
2	Sep 1	Introduction to Linear Regression			
	Sep 3				
3	Sep 8	Generalized Linear Regression	HW1 Out		
	Sep 10			Q1	
4	Sep 15	Sequential Optimization; Regularization			
	Sep 17				Team Formation Due
5	Sep 22	Logistic Regression			
	Sep 24	Generative and Discriminative Classification	HW1 Due HW2 Out	Q2	
6	Sep 29	Decision Trees			

	Oct 1	Random Forests			
7	Oct 6	*Fall Recess*			
	Oct 8	Naive Bayes Classifiers		Q3	Project Proposal Due
8	Oct 13	Clustering and *k*-means			
	Oct 15	SVD and PCA			
9	Oct 20	Gaussian Mixture Models			
	Oct 22		HW2 Due HW3 Out		
10	Oct 27	Hierarchical Clustering			
	Oct 29	DBSCAN Clustering		Q4	
11	Nov 3	Exam Review			
	Nov 5	*Course Practical Exam*			
12	Nov 10	Neural Networks, Backpropagation, MLPs			
	Nov 12				Project Midpoint Report Due
13	Nov 17	CNNs and RNNs			
	Nov 19		HW3 Due		
14	Nov 24	Transformers and Self-Attention		Q5	
	Nov 26	*Thanksgiving Break*			
15	Dec 1	Final Project Presentations			
	Dec 3				
16	Dec 8	Project Workday and Open Discussion			

F	Dec 14	*Final Project Report Due*			Final Project Report Due at 10:50am, the end of our course final exam slot.

Grading Policy:

Deliverables may be turned in until 11:59pm on the due date, after which they are considered +1 day late for every additional 12am-11:59pm period that passes. For every day late, a penalty of 10% off the total score for the deliverable will be subtracted up to 5 days (50% reduction in possible total score). For deliverables that arrive 6 or more days late, a score of 0 will be assigned. Assessments will be comprised of:

Assessment Tool	**% of Grade**
Homeworks (x3)	30%
In-class Quizzes (x5)	10%
In-Class Course Practical Exam	10%
Class Participation	10%
Course Project	40%
TOTAL	**100%**

Letter grades will be assigned in the following buckets based on final grade.

Score Band	**Letter Grade**
90-100	A
80-89	B
70-79	C
60-69	D
0-59	F

Description of Graded Components

Homeworks (3 assignments at 10% each; 30% of final grade):
Through coding homework assignments, students will implement and apply core Machine Learning concepts across supervised and unsupervised learning, demonstrating competency in linear algebra, probability and statistics, optimization, and practical programing implementation. Homework coding assignments that will involve implementing core concepts using frameworks like PyTorch, then training and executing corresponding machine learning models on realistic data. These coding assignments will be graded based on code correctness in terms of producing expected output, as well as through a written report documenting the code design choices and results of the training and evaluation experiments associated with the assignment.

In-Class Quizzes (5 quizzes at 2% each; 10% of final grade total):
In-class quizzes will evaluate students' ability to identify structures, conventions, and algorithmic details underpinning machine learning technologies. Quizzes will occur during a subset of class periods and will involve questions about recently covered topics.

In-Class Course Practical Exam (10% of final grade):
An in-class assessment to be held on November 5 to evaluate understanding of the material covered so far.

Class Participation (10% of final grade):
Students will earn points for class participation both in-person and online. In-person, students earn points by being present and attentive during class when called on to answer questions or work out exercises during class. Online, students earn points through answering questions and generally being helpful on EdStem.

Course Project (40% of final grade):
Students will design and carry out a course project that utilizes machine learning techniques. Students will work individually or in small teams (e.g., 2-3). These research-focused projects will develop machine learning models and analysis of data for an existing problem, or formulate a new problem altogether. Students will leverage tools, concepts, and techniques presented in the class. The project involves identifying a data-centric need that machine learning could resolve, data sources available to inform the problem and method, and the techniques needed to approach it. The grading distribution for deliverables of the course project as well as expectations for each deliverable are detailed below. The deliverables include a project proposal (1-2 pages

single space), a mid project report (4-8 pages single space), final presentation (timed, with time for questions), and a final report (~8 pages single space for the main document, up to 15 with references and appendix). Reports will be graded based on clarity and completeness. The project is total 40% of final grade with the following breakdown:

Project Component	**% of Grade**
Team Formation	1%
Proposal	4%
Midpoint Report	10%
Final Presentation	10%
Final Report	15%
TOTAL	**40%**

Details about each project deliverable follow:

The *project proposal* (about 2 pages) should answer the following questions clearly in a sentence and/or a few paragraphs each, as appropriate: What are you trying to do? Articulate your objectives using absolutely no jargon. How is it done today, and what are the limits of current practice? What is new in your approach and why do you think it will be successful? Who cares? If you are successful, what difference will it make? What are the risks? What could go wrong, and how will you pivot early on if that happens? How much will it cost? That is, what resources will you need in terms of time and computation? Are these reasonable for a semester and what access you have? Identify two milestones along the way to your finished project. (4% of grade total)

The *project midpoint report* (4 pages) should cover these questions in detail: What are you trying to do? Articulate your objectives using absolutely no jargon. Who cares? If you are successful, what difference will it make? Additionally, it should cover related work in detail, as well as document the proposed method and what is new in that approach, as well as how you plan to conduct experiments, what hypotheses these experiments test, and how you will evaluate the quality of the results. If you have preliminary results to report, they should be discussed here. Additionally, you should revisit your milestones from the proposal and document your progress with respect to these, and propose updated milestones for the remainder of the semester. A rubric follows (10% of grade total):

- Conforms to the ICML LaTeX style templates here with a 4 page limit excluding references. (01%)
- Contains an Introduction section that succinctly and clearly answers Heilmeier questions 1–4. (03%)
- Contains a Related Work section that indexes related literature relevant for the proposed work. (01%)
- Contains a Preliminary Methods section outlining the high level hypothesis or claim and at least one testable hypothesis with dependent and independent variables identified alongside the metric by which the dependent variable will be measured. The hypothesis should be stated as a truth condition, and the associated statistical or other concrete evaluation test planned to reject or accept the hypothesis should be included. (02%)
- Contains a Preliminary Results and Next Steps section outlining any experiments conducted so far aligned to the stated hypotheses. Lays out weekly or biweekly milestones up to the final report date and an associated division of responsibilities among team members. (03%)

The *project final presentation* will be given by your team in class and graded on clarity, completeness, and presentation quality. Each member of the group should present for about equal time. Your slides should be numbered so that you can receive detailed feedback from the instructor and to help with guiding the Q&A session after your talk. Presentation and Q&A times will be set based on the number of projects in the course during the semester. Your presentation should convey:

- What your project aimed to do, articulating your objectives using absolutely no jargon.
- What was done before and what your key insight was to go beyond those approaches in your course project.
- Your findings, each presented as: a hypothesis, the methodology you used to evaluate the hypothesis, your experiments and results, and a brief discussion of whether your findings support the hypothesis. (The bulk of your presentation time should be here.)
- What next steps your work leaves open for future research that are not in scope for the course project.

Your final project presentation grade will additionally incorporate your engagement with your classmates' projects. You should aim to engage projects with meaningful, incisive questions. Rubric (10% of final grade):

- Final project presentation clarity (appropriate level of project detail) (02%)

- Final project presentation completeness (containing all required elements) (03%)
- Final project presentation quality (including balancing time between members) (02%)
- Meaningful and incisive question/discussion about another team's project (01.5%)
- Meaningful and incisive question/discussion about another, different team's project (01.5%)

The *project final report* (8 pages) should answer all the following questions in detail, and is expected to include revised content from the midterm report based on feedback your team received at that time. Rubric (15% of final grade):

- Conforms to the ICML style templates linked above with an 8 page limit. The page limit excludes: references, Limitations sections, and optional Appendix. (01%)
- Contains an Introduction section that succinctly and clearly answers Heilmeier questions 1–4. (03%)
- Contains a Related Work section that indexes related literature in the submitted work. This section should contain updates since the midterm report. (03%)
- Contains a Methods and Experiments section outlining the high level hypothesis or claim and at least one testable hypothesis with dependent and independent variables identified alongside the metric by which the dependent variable will be measured. The hypothesis should be stated as a truth condition, and the associated statistical or other concrete evaluation test planned to reject or accept the hypothesis should be included. The experimental protocol for testing these hypotheses should be described clearly in this section. (04%)
- Contains a Results and Conclusions section outlining any experiments conducted aligned to the stated hypotheses and how these results support or contradict those hypotheses. Where applicable, appropriate statistical tests and corrections should be run to make quantified claims. (03%)
- Contains a Future Work section divided into two parts:
    - Research Future Work: what could be built on this work?, similar to what one would see in a published research paper. (00.5%)
    - Team Future Work: what you, the team, are actually considering, if any. If you are not planning to continue the work in any way, state that. (00.5%)

Course Policies

Attendance and/or Participation

Attend class whenever possible.

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. Review Georgia Tech's Honor Code and the student Code of Conduct.

Any student suspected of cheating or plagiarism 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.

Core IMPACTS

Core IMPACTS is the University System of Georgia's General Education curriculum. If you are teaching a course that counts towards Core IMPACTS, you should include a syllabus statement about the Core area and associated career competencies. This resource developed by the Center for Excellence in Teaching and Learning and Online Education at Georgia State University includes template syllabus statements for each of the Core IMPACTS areas that you may adapt for your course.

Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services (404-894-2563) as soon as possible to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail me 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. The Student-Faculty Expectations articulate some basic expectations that you can have of me and that I have of you. 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.

Collaboration, Group Work, and Use of Generative AI

AI-Based Assistance: We are using the AI assistant policy developed by David Joyner and shared by other classes at Georgia Tech (CS 7643 Deep Learning). The summary is that you should treat your AI source like a human source, with all accompanying plagiarism implications:

We treat AI-based assistance, such as ChatGPT and Copilot, the same way we treat collaboration with other people: you are welcome to talk about your ideas and work with other people, both inside and outside the class, as well as with AI-based assistants.

However, all work you submit must be your own. You should never include in your assignment anything that was not written directly by you without proper citation (including quotation marks and in-line citation for direct quotes).

Including anything you did not write in your assignment without proper citation will be treated as an academic misconduct case. If you are unsure where the line is between collaborating with AI and copying AI, we recommend the following heuristics:

Heuristic 1: Never hit "Copy" within your conversation with an AI assistant. You can copy your own work into your own conversation, but do not copy anything from the conversation back into your assignment. Instead, use your interaction with the AI assistant as a learning experience, then let your assignment reflect your improved understanding.

Heuristic 2: Do not have your assignment and the AI agent open at the same time. Similar to the above, use your conversation with the AI as a learning experience, then close the interaction down, open your assignment, and let your assignment reflect your revised knowledge.

This heuristic includes avoiding using AI directly integrated into your composition environment: just as you should not let a classmate write content or code directly into your submission, so also you should avoid using tools that directly add content to your submission.

Deviating from these heuristics does not automatically qualify as academic misconduct; however, following these heuristics essentially guarantees your collaboration will not cross the line into misconduct.

Collaboration and Academic Integrity:

- Plagiarism is a serious offense. You are responsible for completing your own work. You are not allowed to copy and paste, or paraphrase, or submit materials created or published by others, as if you created the materials. All materials submitted must be your own.
- You may discuss high-level ideas with other students at the “whiteboard” level (e.g., how cross validation works, use hashmap instead of array) and review any relevant materials online. However, each student must write up and submit his or her own answers.
- You must not put your code on public domain (e.g., public GitHub), because a (future) student could copy your code. That student obviously violates the honor code, and you may also be implicated.
- All incidents of suspected dishonesty, plagiarism, or violations of the Georgia Tech Honor Code will be subject to the institute’s Academic Integrity procedures (e.g., reported to and directly handled by the Office of Student Integrity (OSI)). Consequences can be severe, e.g., academic probation or dismissal, grade penalties, a 0 grade for assignments concerned, and prohibition from withdrawing from the class.

Extensions, Late Assignments, & Re-Scheduled/Missed Exams

In general, we will not be able to accommodate extensions, late assignments, or re-scheduling of the in-class exam or in-class quizzes. If you know that you will be unable to attend a particular lecture period that includes an evaluation on the schedule, let us know as soon as possible.

Campus Resources for Students

Undergraduate Student Academic Success Resources:

A list of resources for undergraduate students’ academic success and information about advising can be found at Success at Tech.

- Academic Support: Academic Success and Advising (a unit in the Office of Undergraduate Education & Student Success) provides free support for your courses. Students can attend scheduled supplemental review (PLUS) sessions, stop by Drop-In Tutoring, or schedule a one-on-one appointment through Knack. To explore what options work best for you, please visit us online at success.gatech.edu/tutoring, email us at tutoring@gatech.edu, or come see us at Clough Undergraduate Learning Commons, Suite 283.

Graduate Student Academic and Professional Success Resources:

A list of resources for graduate students is given on the Office of Graduate and Postdoctoral Education website. Specific information for current graduate students includes

- Academic Resources such as the Communications Center, Language Institute, Library, Catalog, Registrar, resources for conducting research, Advocacy and Conflict Resolution resources, and how to manage unexpected situations that may impact your academic performance;
- Student Resources such as Campus Services, Child Care/Family programs, Health & Wellness, Career Services, and the Student Resource Guide; and
- *Professional Development such as the programming from the Career Center and other professional development resources and events"*

Student Well-Being:

At Georgia Tech, we are concerned about your overall physical, social, and mental well-being. A comprehensive list of wellness related resources has been compiled and maintained by the Office of the Vice President for Student Engagement and Well-being (student-resource-guide (gatech.edu)