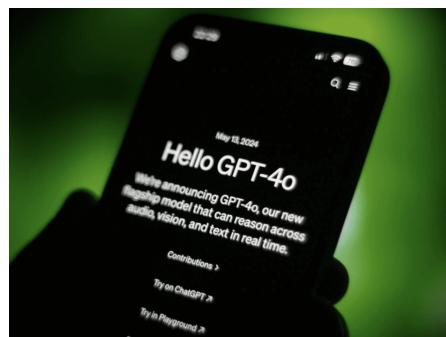
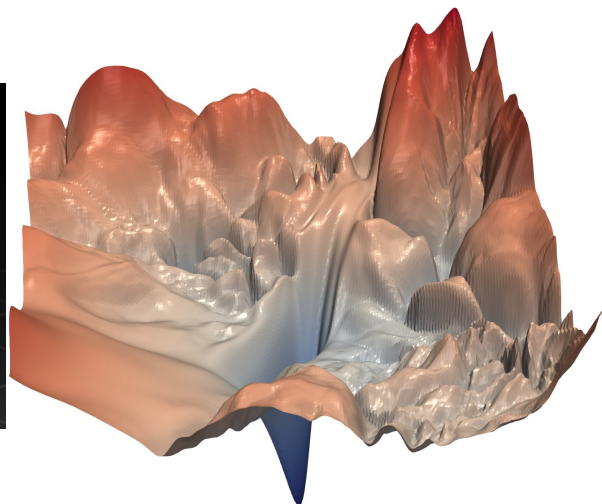
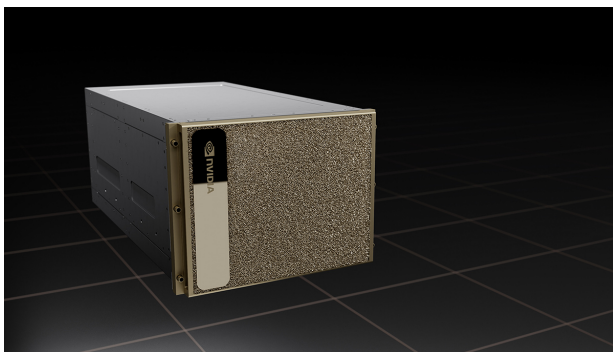


Recent Announcements

CS 5781 Machine Learning Engineering (2024FA)

[Jump to Today](#) Edit

Machine learning is increasingly driven by advances in the underlying hardware and software systems. This course will focus on the challenges inherent to engineering machine learning systems to be correct, robust, and fast. The course walks through the development of a software library for machine learning from scratch, with each assignment requiring students to build models in their own library. Topics will include: tensor languages and auto-differentiation; model debugging, testing, and visualization; fundamentals of GPUs; compression and low-power inference. Guest lectures will cover current topics from ML engineers.



Instructor: [Prof. Alexander \(Sasha\) Rush](http://rush-nlp.com/)  (<http://rush-nlp.com/>)

OH: Monday 3pm (Bloomberg 368)

TAs:

TA OH: Wednesday After Class. Zoom before assignments.

Credits: 3

Course Frequency: Fall Term

Times: Mon - Wed 5:55-7:10pm

Location: Bloomberg Auditorium

Ed Discussions:

MiniTorch: <https://minitorch.github.io/>  [\(https://minitorch.github.io/\)](https://minitorch.github.io/)

Survey:

Course Structure

CS 5781 is a course designed for students interested in the engineering aspects of ML systems. Instead of surveying different tasks and algorithms in ML, the course will focus on the end-to-end process of implementing, optimizing, and deploying a specific model. By limiting ourselves to a fixed model architecture, we will be able to better examine each aspect of the pipeline leading to final deployment, and examine the trade-offs in training, debugging, testing, and deployment, both at a low-level (hardware) and at a high-level (user tools). CS 5781 will be less mathematically demanding than other ML courses, although it does require familiarity with matrices and derivatives. On the other hand, it will be significantly more programming intensive. Each assignment will require completing significant programming exercises in Python, leading up to full implementation of ML systems.

This year the course targets non-linear, dense logistic regression, roughly “deep learning”, models. The following are the main units covered. There will be additional sub-units throughout the semester.

- [Intro \(https://canvas.cornell.edu/courses/67762/modules/395207\)](https://canvas.cornell.edu/courses/67762/modules/395207)
- [Unit 0: Fundamentals \(https://canvas.cornell.edu/courses/67762/modules/395208\)](https://canvas.cornell.edu/courses/67762/modules/395208)
- [Unit 1: Autodifferentiation \(https://canvas.cornell.edu/courses/67762/modules/395209\)](https://canvas.cornell.edu/courses/67762/modules/395209)
- [Unit 2: Tensors \(https://canvas.cornell.edu/courses/67762/modules/395210\)](https://canvas.cornell.edu/courses/67762/modules/395210)
- [Unit 3: Efficiency \(https://canvas.cornell.edu/courses/67762/modules/395211\)](https://canvas.cornell.edu/courses/67762/modules/395211)
- [Unit 4: Networks \(https://canvas.cornell.edu/courses/67762/modules/395212\)](https://canvas.cornell.edu/courses/67762/modules/395212)

Grading

Grading for the course is:

50% Modules HWs (for late submissions: 10% docked per day)

25% Course Midterm

10% Course Attendance (quiz completion)

15% Quiz Grades

Student Outcomes

- Mastery of the key algorithms for training and executing core machine learning methods.
- Understanding of the computational requirements of running these systems.
- Practical ability to debug, optimize, and tune existing models in production environments.
- Skills to develop front-ends to easily interact with and explain predictive systems.
- Understanding how bias can be propagated and magnified by ML systems.
- Facility to compare and contrast different systems along facets such as accuracy, deployment, and robustness.

Preparation

Prerequisites: CS 2110 or equivalent programming experience

Math: Students need to be comfortable with calculus and probability, primarily differentiation and basic discrete distributions. The course does not require proofs or extensive symbolic mathematics.

CS: This course is programming intensive. Students should have strong familiarity with Python and ideally some form of numerical library (e.g. numpy, scipy, scikit-learn, torch, tensorflow). Students should have familiarity with foundational CS concepts such as memory requirements and computational complexity.

Assessment and Deliverables

The assessment structure of MLE is completely problem-set and quiz-based. Throughout the semester there will be 5 problem sets (roughly every two weeks). Students may work in teams, but must submit their own implementations.

The goal of the class is for each student to build their own ML Framework from scratch. Each assignment adds one component to the framework, and by the end of the semester students will be able to efficiently train ML models efficiently with their own framework.

The problem sets:

* Assignment 0: Testing, Modules, and Visualization

* Assignment 1: Auto-Derivatives and Training

* Assignment 2: All about Tensors

* Midterm Exam

* Assignment 3: Speeding Things Up

* Assignment 4: Building Real Models

FAQ

Q: What resources do I need to complete the class?

A: The course will require you to have a python development environment set up, ideally on your own machine or on a cloud server. It does not need to be very powerful nor will that help you do better in the class. We will have some lectures using GPUs, but will use Google Colab for these lectures.

Q: What technologies do I need to know to complete the class?

A: This is a software engineering style course, and so we recommend that you have a strong background in standard tools such as Git and GitHub, Python, and command-line programming.

Q: What math do I need to know to complete the class?

A: This course will require light-undergraduate level calculus and vector manipulation. We will provide resources for reviewing these aspects in homework assignments.

Q: How will the course schedule interact with Project Studio?

There will be three Thursday lectures which will be moved to Sunday due to interaction with Project Studio **Maker Days**.

Methods of Assessing Student Achievement

The assignments for this class primarily consist of completing programming assignments developing ML systems. There will be a few written questions asking students to explain reasoning and plan for engineering aspects. These will be graded for accuracy and completeness. Most points however will be assigned based on the accuracy of code implementation as assessed by automated tests and predictive accuracy of final systems.

Exams will consist of multiple choice questions based on code examples and in-class problems, as well as long-form written problems which will take the form of word problems asking students to assess and diagnose real-world systems.

Topics Covered

- Implementation of backpropagation engines
- Low-level interaction with modern hardware GPU
- High-level autodifferentiation in modern libraries
- Core models for vision, text, and recommendations
- Debugging large scale machine learning systems
- Visualization, hyperparameter tuning, and deployment
- Interactions with data processing systems

Student Outcomes

- Mastery of the key algorithms for training and executing core machine learning methods.
- Understanding of the computational requirements of running these systems.
- Practical ability to debug, optimize, and tune existing models in production environments.
- Skills to develop frontends to easily interact with and explain predictive systems.
- Understanding how bias can be propagated and magnified by ML systems.
- Facility to compare and contrast different systems along facets such as accuracy, deployment, and robustness.

Academic Integrity

Each student in this course is expected to abide by the Cornell University Code of Academic Integrity. Any work submitted by a student in this course for academic credit will be the student's own work. The policy can be found on the university's website here:

<https://theuniversityfaculty.cornell.edu/academic-integrity/>

[\(https://theuniversityfaculty.cornell.edu/academic-integrity/\)](https://theuniversityfaculty.cornell.edu/academic-integrity/).

[Academic Integrity \(https://theuniversityfaculty.cornell.edu/academic-integrity/\)](https://theuniversityfaculty.cornell.edu/academic-integrity/)

You are encouraged to study together and to discuss information and concepts covered in lecture and the sections with other students. You can give "consulting" help to or receive "consulting" help from such students. However, this permissible cooperation should never involve one student having possession of a copy of all or part of work done by someone else, in the form of an e-mail, an e-mail attachment file, a diskette, or a hard copy.

Should copying occur, both the student who copied work from another student and the student who gave material to be copied will both automatically receive a zero for the assignment. Penalty for violation of

this Code can also be extended to include failure of the course and University disciplinary action.

During examinations, you must do your own work. Talking or discussion is not permitted during the examinations, nor may you compare papers, copy from others, or collaborate in any way. Any collaborative behavior during the examinations will result in failure of the exam, and may lead to failure of the course and University disciplinary action.

Generative AI

Students will be required to abide by the generative AI policy specified in each of the assignment releases. Use of unauthorized tools or resources will be handled by the discretion of the course staff with the possibility of failing or low-grades on assignments.

COVID-19 Guidelines

The course will follow Cornell and Cornell Tech COVID-19 guidelines.

Remote participation is not permitted. Masks are required in class for all in attendance, regardless of vaccination status.

Students with Disabilities

Your access in this course is important. Please give me your Student Disability Services (SDS) accommodation letter early in the semester so that we have adequate time to arrange your approved academic accommodations. If you need an immediate accommodation for equal access, please speak with me after class or send an email message to me and/or SDS at sds_cu@cornell.edu (mailto:sds_cu@cornell.edu). If the need arises for additional accommodations during the semester, please contact SDS. You may also feel free to speak with Student Services at Cornell Tech who will connect you with the university SDS office.

Religious Observances









Cornell University is committed to supporting students who wish to practice their religious beliefs. Students are advised to discuss religious absences with their instructors well in advance of the religious holiday so that arrangements for making up work can be resolved before the absence.












Statement about our supportive community












The Cornell Tech community is a diverse and vibrant group of students, faculty, and staff. We take our responsibility to look out for one another seriously. As members of this community, your openness and













proactive communication will allow us all to better care for students and respond to their needs, whether they be interpersonal or academic. Please help us continue to build and strengthen our community by reaching out if you are having an issue or are concerned about a fellow student. Contact studentwellness@tech.cornell.edu (mailto:studentwellness@tech.cornell.edu) with concerns and we will make sure to care for one another. In the event of an emergency, please call 911 and Cornell Tech Safety Security at 646-971-3611 (This number is also located on the back of your Cornell ID), when safe to do so.















Course Summary:















Date	Details	Due
Mon Aug 26, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057064&include_contexts=course_67762	5:55pm to 7:10pm
	 CS 5781 Machine Learning Engineering (2023FA) https://canvas.cornell.edu/calendar?event_id=1057067&include_contexts=course_67762	6pm to 7:30pm
Wed Aug 28, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057052&include_contexts=course_67762	5:55pm to 7:10pm
	 CS 5781 Machine Learning Engineering (2023FA) https://canvas.cornell.edu/calendar?event_id=1057068&include_contexts=course_67762	6pm to 7pm
Mon Sep 2, 2024	 [No Class] https://canvas.cornell.edu/calendar?event_id=1057056&include_contexts=course_67762	12am
Wed Sep 4, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057035&include_contexts=course_67762	5:55pm to 7:10pm
Thu Sep 5, 2024	 Module 0 https://canvas.cornell.edu/courses/67762/assignments/679320	due by 11:59pm
Mon Sep 9, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057035&include_contexts=course_67762	5:55pm to 7:10pm












Date	Details	Due
	event_id=1057034&include_contexts=course_67762)	
Wed Sep 11, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057044&include_contexts=course_67762)	5:55pm to 7:10pm
Mon Sep 16, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057037&include_contexts=course_67762)	5:55pm to 7:10pm
Wed Sep 18, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057045&include_contexts=course_67762)	5:55pm to 7:10pm
Mon Sep 23, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057038&include_contexts=course_67762)	5:55pm to 7:10pm
Wed Sep 25, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057046&include_contexts=course_67762)	5:55pm to 7:10pm
Mon Sep 30, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057039&include_contexts=course_67762)	5:55pm to 7:10pm
	 Module 1 https://canvas.cornell.edu/courses/67762/assignments/679321)	due by 11:59pm
Wed Oct 2, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057047&include_contexts=course_67762)	5:55pm to 7:10pm
Mon Oct 7, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057040&include_contexts=course_67762)	5:55pm to 7:10pm
Wed Oct 9, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057036&include_contexts=course_67762)	5:55pm to 7:10pm
Mon Oct 14, 2024	 [No Class] https://canvas.cornell.edu/calendar?	12am

Date	Details	Due
	event_id=1057057&include_contexts=course_67762)	
Wed Oct 16, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057053&include_contexts=course_67762	5:55pm to 7:10pm
Mon Oct 21, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057041&include_contexts=course_67762	5:55pm to 7:10pm
	 Module 2 https://canvas.cornell.edu/courses/67762/assignments/679326	due by 11:59pm
Wed Oct 23, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057054&include_contexts=course_67762	5:55pm to 7:10pm
Mon Oct 28, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057059&include_contexts=course_67762	5:55pm to 7:10pm
Wed Oct 30, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057055&include_contexts=course_67762	5:55pm to 7:10pm
	 Midterm https://canvas.cornell.edu/courses/67762/assignments/679319	due by 11:59pm
Mon Nov 4, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057060&include_contexts=course_67762	5:55pm to 7:10pm
Wed Nov 6, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057048&include_contexts=course_67762	5:55pm to 7:10pm
Mon Nov 11, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057061&include_contexts=course_67762	5:55pm to 7:10pm
Wed Nov 13, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057049&include_contexts=course_67762	5:55pm to 7:10pm

Date	Details	Due
Mon Nov 18, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057062&include_contexts=course_67762	5:55pm to 7:10pm
Wed Nov 20, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057050&include_contexts=course_67762	5:55pm to 7:10pm
Mon Nov 25, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057042&include_contexts=course_67762	5:55pm to 7:10pm
Tue Nov 26, 2024	 Module 3 https://canvas.cornell.edu/courses/67762/assignments/679335	due by 11:59pm
Wed Nov 27, 2024	 [No Class] https://canvas.cornell.edu/calendar?event_id=1057063&include_contexts=course_67762	12am
Mon Dec 2, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057043&include_contexts=course_67762	5:55pm to 7:10pm
Wed Dec 4, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057051&include_contexts=course_67762	5:55pm to 7:10pm
Mon Dec 9, 2024	 Lecture https://canvas.cornell.edu/calendar?event_id=1057058&include_contexts=course_67762	5:55pm to 7:10pm
	 Module 4 https://canvas.cornell.edu/courses/67762/assignments/679343	due by 11:59pm
Tue Dec 10, 2024	 Final Assignment https://canvas.cornell.edu/courses/67762/assignments/679315	due by 11:59pm
	 Corrections https://canvas.cornell.edu/courses/67762/assignments/679311	
	 Extra Credit - Autodiff Puzzles https://canvas.cornell.edu/courses/67762/assignments/679313	

Date	Details	Due
	 <u>Extra Credit - Tensor Puzzles</u> (https://canvas.cornell.edu/courses/67762/assignments/679314)	
	 <u>Lecture 14</u> (https://canvas.cornell.edu/courses/67762/assignments/679316)	
	 <u>Lecture 15</u> (https://canvas.cornell.edu/courses/67762/assignments/679317)	
	 <u>Lecture 4</u> (https://canvas.cornell.edu/courses/67762/assignments/679318)	
	 <u>Module 0.2</u> (https://canvas.cornell.edu/courses/67762/assignments/679310)	
	 <u>Module 1.0</u> (https://canvas.cornell.edu/courses/67762/assignments/679322)	
	 <u>Module 1.1</u> (https://canvas.cornell.edu/courses/67762/assignments/679323)	
	 <u>Module 1.2</u> (https://canvas.cornell.edu/courses/67762/assignments/679324)	
	 <u>Module 1.3</u> (https://canvas.cornell.edu/courses/67762/assignments/679325)	
	 <u>Module 2.0a</u> (https://canvas.cornell.edu/courses/67762/assignments/679327)	
	 <u>Module 2.0b</u> (https://canvas.cornell.edu/courses/67762/assignments/679328)	
	 <u>Module 2.1</u> (https://canvas.cornell.edu/courses/67762/assignments/679329)	
	 <u>Module 2.2a</u> (https://canvas.cornell.edu/courses/67762/assignments/679330)	
	 <u>Module 2.2b</u> (https://canvas.cornell.edu/courses/67762/assignments/679331)	

Date	Details	Due
	 <u>Module 2.3</u> (https://canvas.cornell.edu/courses/67762/assignments/679332)	
	 <u>Module 2.4</u> (https://canvas.cornell.edu/courses/67762/assignments/679333)	
	 <u>Module 2.5</u> (https://canvas.cornell.edu/courses/67762/assignments/679334)	
	 <u>Module 3.0</u> (https://canvas.cornell.edu/courses/67762/assignments/679336)	
	 <u>Module 3.0a</u> (https://canvas.cornell.edu/courses/67762/assignments/679337)	
	 <u>Module 3.1a</u> (https://canvas.cornell.edu/courses/67762/assignments/679338)	
	 <u>Module 3.2a</u> (https://canvas.cornell.edu/courses/67762/assignments/679339)	
	 <u>Module 3.2b</u> (https://canvas.cornell.edu/courses/67762/assignments/679340)	
	 <u>Module 3.3</u> (https://canvas.cornell.edu/courses/67762/assignments/679341)	
	 <u>Module 3.5</u> (https://canvas.cornell.edu/courses/67762/assignments/679342)	
	 <u>Module 4.0</u> (https://canvas.cornell.edu/courses/67762/assignments/679344)	
	 <u>Module 4.1</u> (https://canvas.cornell.edu/courses/67762/assignments/679345)	
	 <u>Module 4.2</u> (https://canvas.cornell.edu/courses/67762/assignments/679346)	
	 <u>Module 4.3</u> (https://canvas.cornell.edu/courses/67762/assignments/679347)	

Date	Details	Due
	 <u>Module 4.4a</u> (https://canvas.cornell.edu/courses/67762/assignments/679348)	
	 <u>Module 4.4b</u> (https://canvas.cornell.edu/courses/67762/assignments/679349)	
	 <u>Week 1 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679350)	
	 <u>Week 2 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679351)	
	 <u>Week 3 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679352)	
	 <u>Week 4 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679353)	
	 <u>Week 5 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679354)	
	 <u>Week 6 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679355)	
	 <u>Week 7 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679356)	
	 <u>Week 8 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679357)	
	 <u>Week 9 Lecture Quizzes</u> (https://canvas.cornell.edu/courses/67762/assignments/679358)	