



CIS 4190/5190

Applied Machine Learning

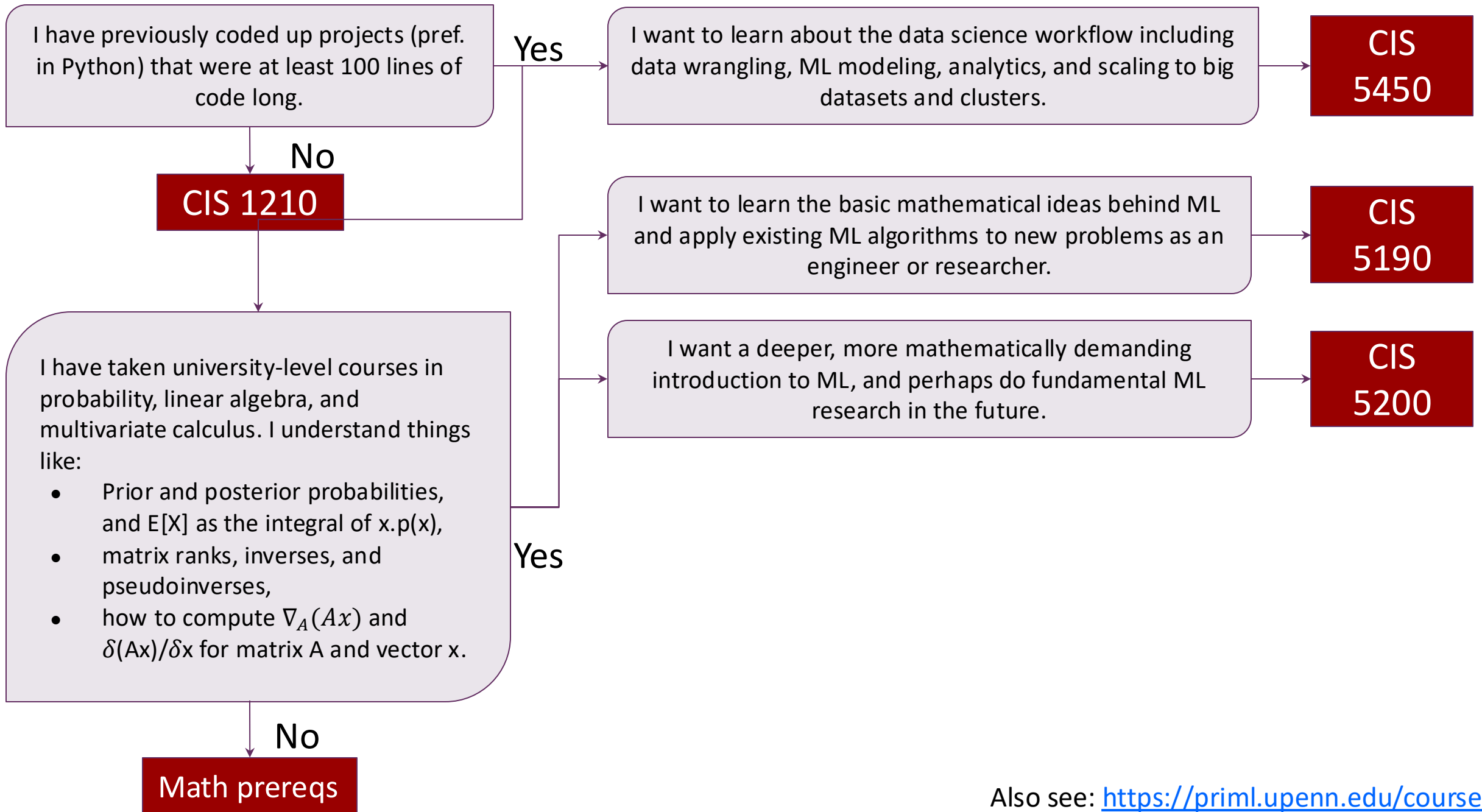
Instructors: Dinesh Jayaraman and Mingmin Zhao

Course Objectives

- **The aim of this course is to produce strong, versatile, and ethical practitioners of machine learning (ML).** Students should have the skills to:
 - Identify opportunities for applying various ML algorithm classes
 - Code and debug their own ML applications
 - Appreciate the ML workflow from data gathering to deployment
- Lectures will focus largely on developing mathematical understanding.
- Assignments will focus on applying this understanding to implement ML solutions.

Prerequisites

- **Math:** University-level courses in probability, linear algebra, and multivariable calculus
 - Understand prior and posterior probabilities, $\mathbb{E}[X] = \int p(x)dx$, etc.
 - Understand matrix ranks, inverses, and eigenvalues
 - Understand how to compute $\nabla_A(Ax)$ for a matrix A and vector x
 - Tested in HW 0
- **Programming:** Previously coded up projects (preferably in Python) that were at least 100 lines of code long
 - We will provide Python help (primer + office hours) for students who know how to program in languages other than Python
 - If you are not confident of your coding skills in any language at all, homework for this class could get very hard!
- Tested partially in HW0 (more on this soon)
- Primers: <https://www.seas.upenn.edu/~cis5190/fall2024/resources.html>



Also see: <https://priml.upenn.edu/courses/>

What Does “Applied” Mean?

- Theoretical ML deals with theoretical analysis, i.e., proving theorems about things like “how much data you need in order to have a 99% chance of achieving 95% accuracy using <your favorite> ML algorithm”?
- Applied means that we will not emphasize these aspects.
 - ✧ **You will not be asked to prove theorems.**
 - We might cover them in class on occasion when it is pedagogically useful to explain the practice of ML, but we won’t ask you to reproduce a theorem in exams.
 - ✧ **But you will be solving math problems** (number and algebra) with pen and paper
 - PS: This is also the kind of thing you might get asked in an interview for ML engineer roles in industry that are of interest to many students taking this class.

Humans of CIS 4190/5190 Fall'24

Instructor Introductions

Dinesh Jayaraman

Assistant Professor, CIS

GRASP laboratory

<https://www.seas.upenn.edu/~dineshj/>



Mingmin Zhao

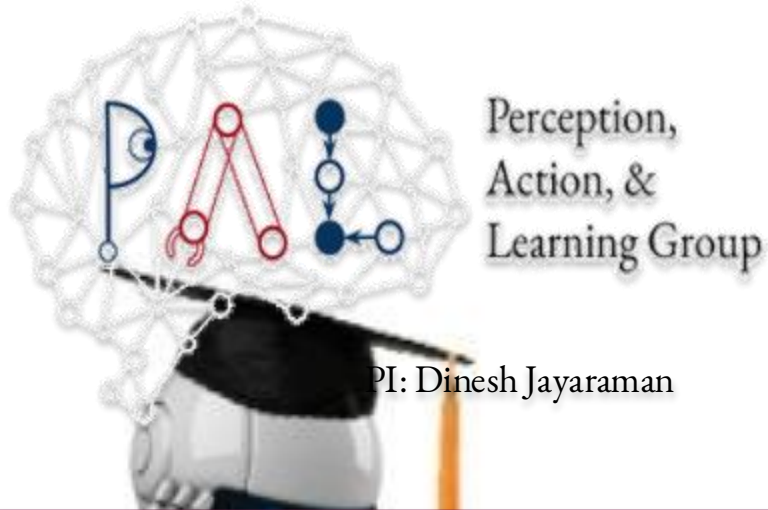
Assistant Professor, CIS

<https://www.cis.upenn.edu/~mingminz/>



Research Areas – Dinesh: Robot Learning

Some recent highlights:



Fast, Flexible, and Simple Robot Teaching
By Focusing on Task-Relevant Information
In Observations, Representations, Rewards, & Data

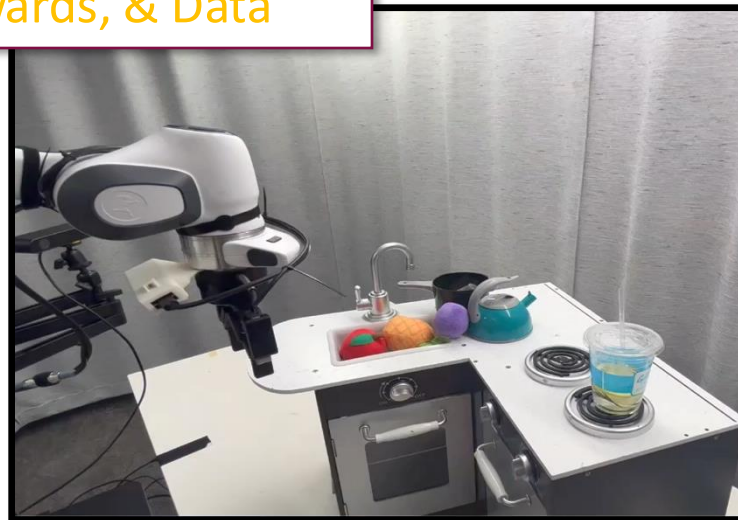
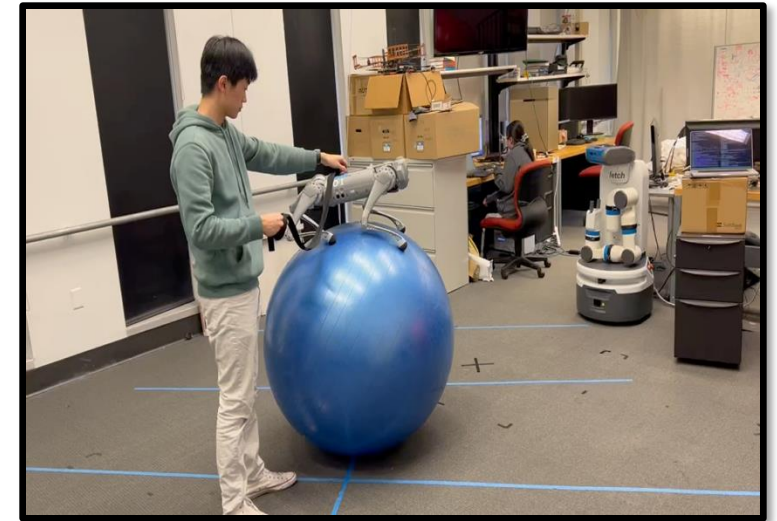
Algorithms validated across many robot platforms and tasks:

- Manipulation & mobile manipulation,
- Quadruped locomotion,
- Dexterous hands,
- Quadrotor flight, ..

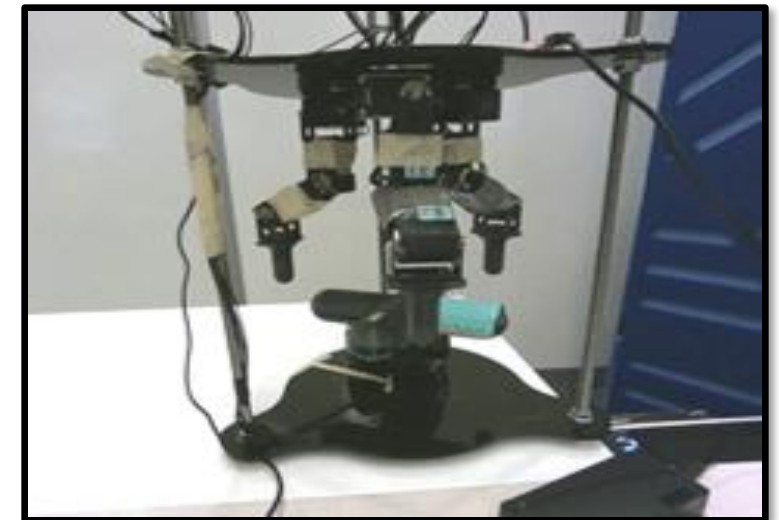
folding a napkin 3 times over



walking on a yoga ball

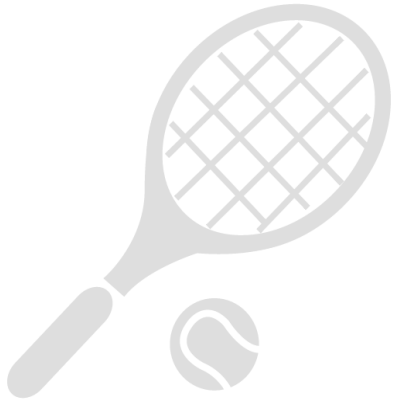


following a 5-step cooking recipe



tightening a screw (and checking)

Hobbies



Research Areas - Mingmin: Multimodal Learning and Sensing

In addition to:

Vision



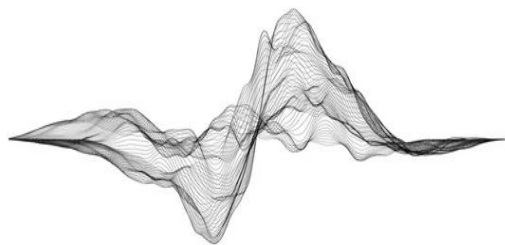
Text



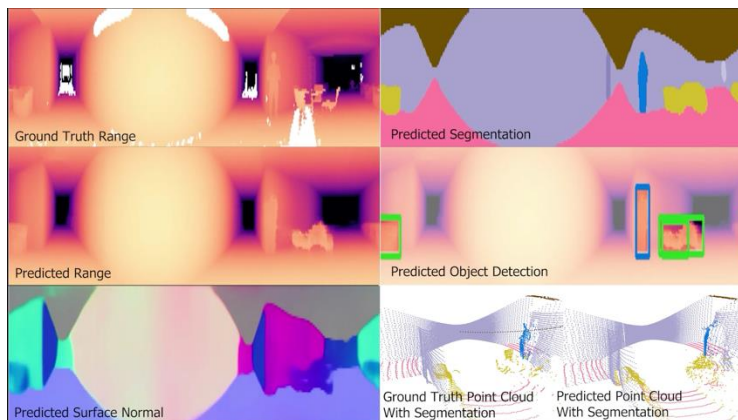
We also look at:

Radio/Wireless

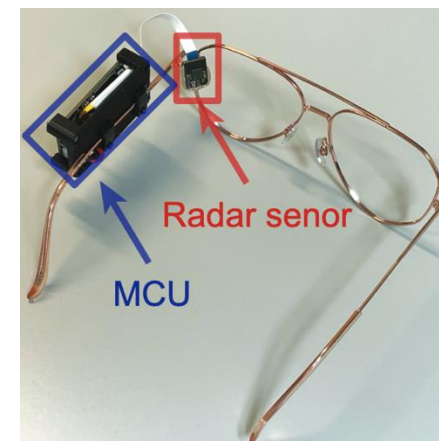
Acoustic/Ultrasound



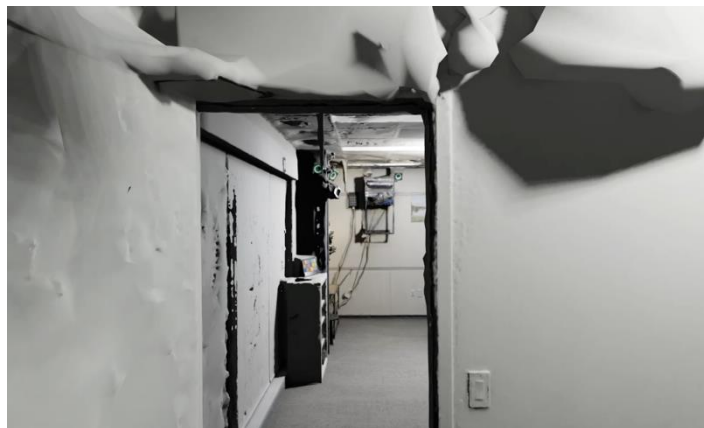
See Through Occlusions & X-Ray Vision



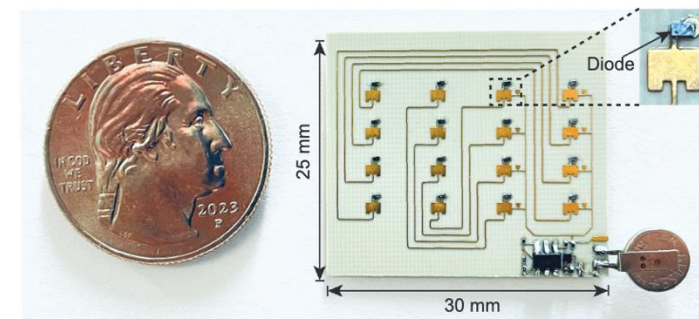
Smart Eyeglasses for Ocular Health &
Cognitive State Tracking



Sound Rendering / Acoustic Modeling



Sub-mm and NLOS Motion Capture



Life & Hobbies

Moko



Moki



Who We Are: The Course Team Of Fall '24



Hersh



Frank
(Tongtong)



Deeksha



Zelong



Guangyao



Chandler



Aurora
(Jianing)



Jocelyn



Daniel



Edward



Haorui



Helen

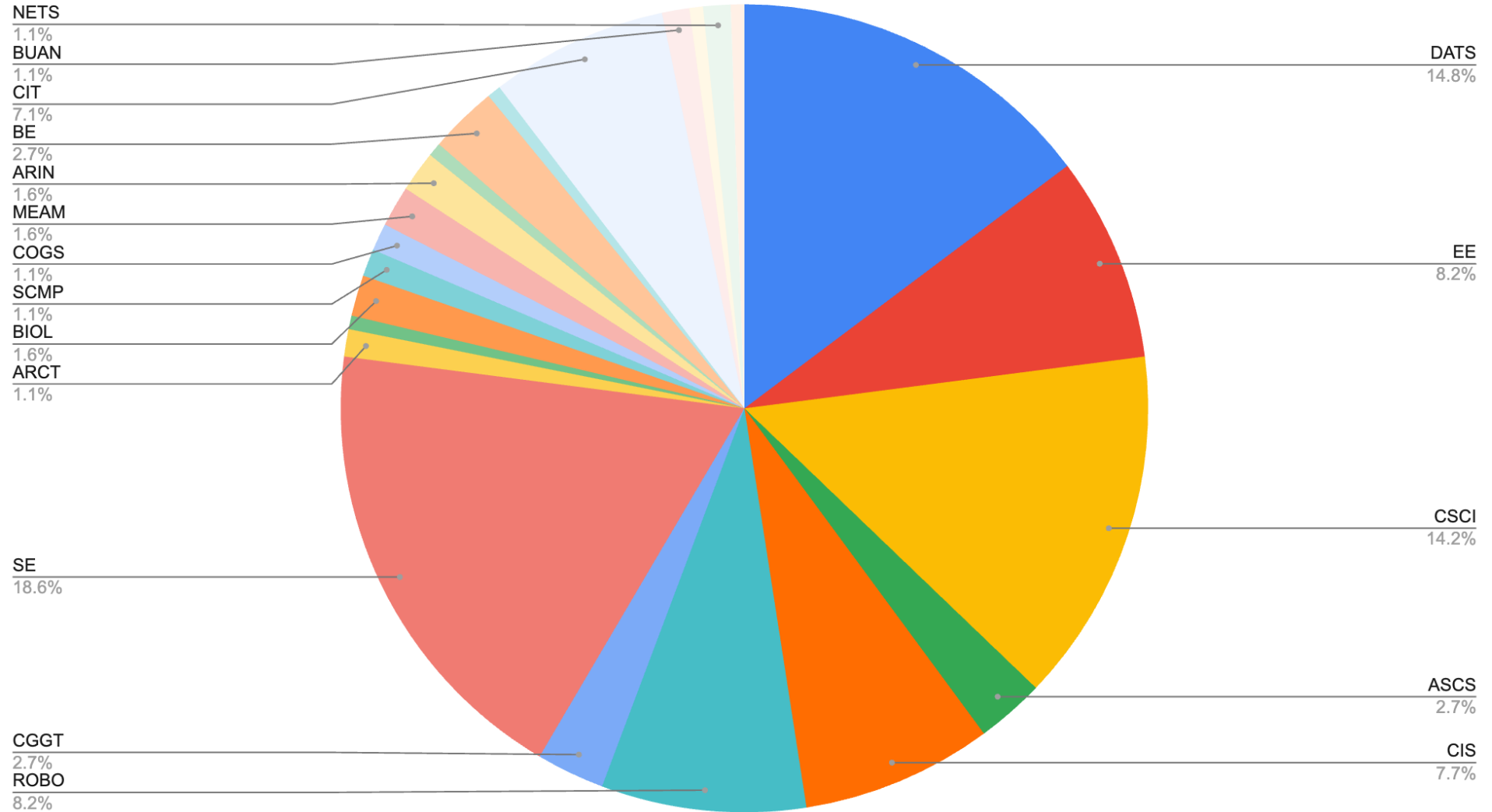


Wendy

Get to know your TAs too! Slide deck on the course website.

Who You Are: Students Of Fall '24

Count of Primary Major



Waitlists

- Students not yet in the class:
 - We're at classroom capacity (207), and we'll add more students as some students drop out
 - Prioritized by
 - All prereqs satisfied
 - Correct / approximately correct answers to all three problems
 - Program (whether 5190 is a core class) and Year
 - Date of request
 - If issued permits have not been used for 48 hours, we will assume you are not interested, and revoke permits. Drop us a note if you would like us to keep your permit valid longer.

Course Schedule

Fall 2024 Course Content and Tentative Schedule

Week	Date*	Modules		
1	Aug 28	Intro		
2	Sep 04	Linear Regression ...		HW0 (Aug 28-Sep 11) 5%
3	Sep 09	... Linear Regression		
4	Sep 16	Logistic Regression		HW1 (Sep 11-Oct 02) 5%
5	Sep 23	Basics of Neural Networks		
6	Sep 30	KNN and Decision Trees		
7	Oct 7	Ensembles and Clustering		HW2 (Oct 02-Oct 23) 5%
8	Oct 14	Review and Mid-term 1 (Oct 16) (15%)		
9	Oct 21	ML with Images: Convolutional Neural Nets		
10	Oct 28	ML with Text Sequences: RNNs and Transformers		HW3 (Oct 23-Nov 06) 5%
11	Nov 04	ML for Sequential Decision Making: Reinforcement Learning		
12	Nov 11	ML for Recommendations		HW4 (Nov 06-Nov 20) 5%
13	Nov 18	Overview of Generative AI for Images and Text		
14	Nov 25	Ethics		
15	Dec 2	Reviews		
16	Dec 9	Mid-term 2 (Dec 9) (15%)		

Grading Scheme

- 5x Homeworks: 25%
- Project (teams of 3): 30% -- **more details in the coming weeks**
- Mid-term 1: 15%
- Mid-term 2: 15%
- Weekly quizzes: 10% (Full points for 50% correct, on time)
- Good citizenship points: 5% (everyone has them by default, you can lose them for bad behavior e.g. disruptive in class / on Ed discussions)
- Bonus points for extra-good citizenship, e.g. answering lots of questions or posting course-relevant pointers on Ed STEM, bug / typo reporting (more info on this soon): adding up to 5% of course grade

Letter Grading Scheme

- A+: 90+
- A : 85-90
- A-: 80-85
- B+: 75-80
- B : 70-75
- B-: 65-70
- Lower passing grades: 50-65

(Curved, but only upwards. i.e. your grade will only improve.)

Late Policy

- 0.5% of points for any assignment, lost per late hour (rounded up). Max 48 late hours per assignment.
 - e.g. if you submit HW1 20 hours late, you lose $20 * 0.5 = 10\%$ of HW1 points.
 - HW1 carries 5% of course grade => so if you submit 20 hours late, you lose 0.5% of course grade.
- If you have medical reasons to need extensions, send both professors a copy of your medical visit report, and we will grant a 2-day no-penalty extension.
 - Other reasons on a case-by-case basis.

4190 vs 5190

- 5190 will have extra mandatory HW components for 5190, that will serve as optional components for 4190.
- For example:
 - A HW Pset might have 50 points for 5190, of which 5 points might be 5190-only.
 - Then, a student in 4190 will get 100% of the points for that Pset if she scores 45 or higher. i.e. she could skip the 519-only portion and still score 100%.
 - You cannot score more than 100% using optional components. They are not *bonus* points.
 - Separate for Pset and Coding HW: You can't make up for pset points with coding and vice versa.
- The two sections are otherwise identical.

Up-To-Date Information on Course Website

- <https://www.seas.upenn.edu/~cis5190/fall2024/>



Class Links

- Canvas: <https://canvas.upenn.edu/courses/1814284/>
 - Official LMS for Penn courses, but we will mainly rely on:
 - Class website: <https://www.seas.upenn.edu/~cis5190/fall2024/>
 - Lectures and other materials
 - Gradescope: <https://www.gradescope.com/courses/840837/>
 - Weekly quizzes and HWs
 - Ed Discussion: <https://edstem.org/us/courses/64554/discussion/>

Online Communication: Ed Discussion

- We will use **Ed Discussion** for questions, course discussions, announcements
 - Send a message to “instructors” to contact professors and TAs
 - You can contact the professors directly on Ed Discussion or by email (posted on course website); always contact both of us
 - Up to 2 days to respond.
- Always be courteous and respectful of others.
 - Note: “Anonymous” posts on Ed are not anonymous to the course staff
- For confidential feedback, visible to instructors only, use the following Google form. <https://forms.gle/FaGoj4vK9kfbymS6>
 - If you'd like us to follow up with you, you are welcome to leave your information.) If you have ideas to improve the course, better tools to use etc., do not hesitate to tell us!

Office Hours

- Instructors will have 1 weekly OH, at their offices.
 - Dinesh Jayaraman: Levine Hall 401
 - Mingmin Zhao: Levine Hall 503
- TAs will host ~10-15 hours of OH each week.
- Times and venues will be announced soon.

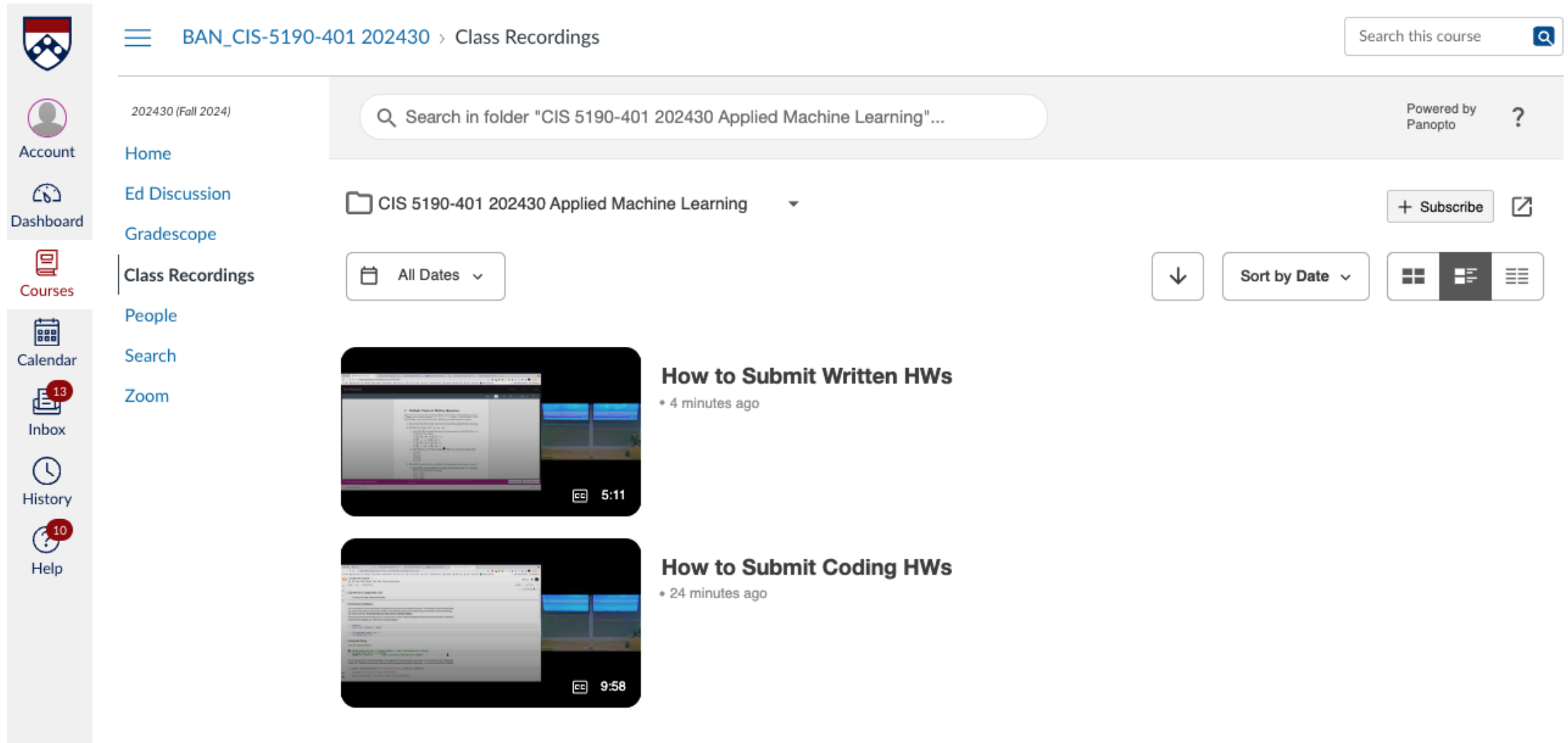
Assignments (+Misc. Course Policy)

Homeworks

- Written Problem Sets:
 - LaTeX encouraged. Handwritten + scanned, at your own risk.
 - Won't be graded if you don't **annotate pages (which question answered on which page) on Gradescope**.
- Coding Homework: Autograder + Gradescope submission of notebook
 - Colab (ipython) notebook skeletons
 - Autograder integrated into notebooks.
- Bonus points up to 20% of the coding HW grade, if you are first to report a bug on Ed (despite our best efforts to avoid them). Incentive to start early!
- Discussion permitted for clarifying the problem specification, but do not ever share solutions / code. Acknowledge all your discussions at the beginning of your report.
 - E.g. "I spoke with TA __ who helped identify the following bug in my code", or "I spoke with classmate __ to better understand questions 3 and 5, but we did not discuss solutions".

Homework Submission Tutorials

- Videos on Canvas > Class Recordings (Note: recordings from a past iteration, so the interfaces you see might be marginally different from what the videos show, don't be alarmed)



The screenshot displays the Canvas LMS interface for a course. On the left is a navigation sidebar with icons for Account, Dashboard, Courses, Calendar, Inbox (13), History, and Help (10). The main content area shows the course path: BAN_CIS-5190-401 202430 > Class Recordings. A search bar at the top right allows searching within the course. Below the course path, there is a search bar for the folder "CIS 5190-401 202430 Applied Machine Learning". The folder name is displayed with a dropdown arrow. To the right of the folder name are a "+ Subscribe" button and a share icon. Below the folder name, there is a date filter set to "All Dates" and a "Sort by Date" dropdown menu. The video list shows two recordings: "How to Submit Written HWs" (5:11, 4 minutes ago) and "How to Submit Coding HWs" (9:58, 24 minutes ago). Each video thumbnail includes a CC icon and a duration indicator.

202430 (Fall 2024)

Search this course

Search in folder "CIS 5190-401 202430 Applied Machine Learning"...

Powered by Panopto ?

CIS 5190-401 202430 Applied Machine Learning

+ Subscribe

All Dates

Sort by Date

How to Submit Written HWs
• 4 minutes ago
5:11

How to Submit Coding HWs
• 24 minutes ago
9:58

Quiz Schedule

- **Weekly Quizzes**


- Expected to take ~5-10 mins per week
 - *Much* easier questions than HW or exams
 - Released every Wednesday
 - Due Thursday 8 days later
 - No late submissions
 - You can miss up to 3 weekly quizzes, with no penalty in your course grade.
- For any quiz, $\geq 50\%$ score \Rightarrow full points towards course grade.
 - You can retake each quiz any number of times, but not **after** the deadline
 - Checks that you are at least loosely following material covered the previous week

First Assignments

- **HW 0: Released already today, due 9/11 at 7.59 p.m. ET**
 - No office hours planned for HW 0
 - You can ask clarification questions via Ed Discussion
 - (only for HW0, not true for other HWs) 50% points on each part (coding & written) = full credit for that part
 - Expected time: ~3 hours.
 - Opportunity for you (and us) to get used to the workflow.

- **Quiz 1: Will release 9/04, due 9/12**

Gradescope

 < ☰

CIS-5190-401 202430
Applied Machine Learning

- Dashboard
- Assignments
- Roster
- Extensions
- Course Settings

Instructors


- Dinesh Jayaraman
- Tongtong Liu
- mingminz

Course Actions

- Unenroll From Course

CIS-5190-401 202430 | Fall 2024

Course ID: 840837

 No Published Grades

Description

Edit your course description on the [Course Settings](#) page.

Things To Do

 Review and publish grades for [HWO Coding](#) now that you're all done grading.

Active Assignments	Released	Due (EDT)	Submissions	% Graded	Published	Regrades
HWO Written Problem Set	AUG 28, 2024 10:00 AM	SEP 11, 2024 7:59 PM Late Due Date: SEP 13, 2024 7:59 PM	1	0%	<input type="radio"/>	ON
HWO Coding	AUG 28, 2024 10:00 AM	SEP 11, 2024 7:59 PM Late Due Date: SEP 13, 2024 7:59 PM	1	100%	<input type="radio"/>	ON

Planned Homework Schedule

- All HWs released on Wednesday, approximately once in 2-3 weeks
- Due on date of release of next HW at 7.59 p.m. ET
- Late policy as announced before: 0.5% points lost for every late hour on each HW part separately, up to a max of 48 hours for each part.

AI / LLM Policy

- Modern AI tools can be of great help in understanding concepts, and we have no concerns about you using ChatGPT, Claude, Gemini, etc. to get alternative explanations for topics. **Always dig a bit deeper when learning from an LLM**, like following up links to make sure the LLM did not simply make something up.
- Given that we are trying to teach general, reusable skills, we require you to **write your homework solutions *without* help from an LLM or from a classmate**. Please note that the exams will be tailored with this in mind (i.e., closed-book exams focusing on the ability to tackle problems) so you should make sure you can solve problems on your own!

Lecture Video Recordings & Slides Policy

- Slides will usually be up ***after* class, not before.**
 - We are often making slides until a few mins before class begins.
 - Release on class website
 - If you really want slides to take notes on, you could consider visiting the websites from past semesters (e.g. search for CIS 5190 Spring 2023 website), but you will have to figure out what changed and what didn't.
- We will video-record lectures but **only release them 1 week after class.**
 - To incentivize in-class participation
 - Exception (immediate video recording release) for weeks 1 and 2 to allow for add/drop and students that enter the class via the waitlist.
 - Released on Canvas > Class Recordings

And Now On To Learning About ML!