Introduction to Machine Learning (Honors)

CS 375(H)
Credits: 3
Mode: Face-to-Face
Location: Central King Building 317
Lecture: Wednesday/Friday, 1:00-2:20pm

Instructor: Kieran Murphy
Email: kieran.murphy@njit.edu
Office: 4315 GITC
Office Hours: Fridays 2:30-3:30pm. If my posted office hours don't work for you, just email me -- I'm happy to find another time.
Note: I will make every effort to respond to messages by the end of the following business day. If I haven't responded, please ping me again.

Tutoring. NJIT provides a tutoring service that you may find helpful. Please check the website for updates regularly as they may change the information.

Welcome to machine learning! I'm excited for our journey covering methods that we use to make models of the world through data and clever computational algorithms. Our semester will progress along an axis that roughly tracks the number of learnable parameters, starting at (essentially) zero with k-nearest neighbors and approaching trillions in the largest language models out today. As we progress, we'll rely on insights from statistical learning to guide our utilization of data.

Course Description

Machine learning is the science and engineering of finding useful mappings from data. Over decades, the field has developed a toolkit for choosing what mappings to consider, fitting them with finite data, and determining whether they work beyond the examples seen during training. This course studies that foundation, not as history, but because the same ideas underlie modern AI. By the end, you should be able to look at a new ML system or AI development and recognize many of the ingredients, understand why they are there, and have the foundation to work with them yourself.

Prerequisites

CS 100 and (CS 113 or CS 115) and (MATH 333 or MATH 341) with a grade C or better. Some background in calculus, linear algebra, probability, and Python programming is assumed.

Helpful math resource: Mathematics for Deep Learning chapter (from Dive into Deep Learning by Zhang, Lipton, Li and Smola)
Helpful python resources: A whirlwind tour of python by Jake VanderPlas, NumPy: the absolute basics for beginners, A visual intro to NumPy and data representation by Jay Alammar

Course Textbooks and Resources

Primary Resource: An Introduction to Statistical Learning with Applications in Python (free pdf available through https://www.statlearning.com/)
by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani

Learning Outcomes

By the end of the course, students will be able to:

a. **Formulate a problem as a machine-learning problem**, specifying the data, inputs and outputs, model class, training objective, and criteria for success.
b. **Decompose a machine-learning system into its main ingredients** -- representation, model, loss, optimization, regularization, and evaluation -- and explain how those ingredients interact.
c. **Translate between intuitive, mathematical, and computational descriptions of machine learning**, and use each to explain model behavior and design choices.
d. **Implement, train, and adapt representative machine-learning models**, from classical methods through neural networks, using appropriate computational tools.
e. **Design and interpret empirical evaluations of ML systems**, selecting suitable baselines, data splits, metrics, and diagnostic experiments while recognizing common sources of misleading results.
f. **Analyze an unfamiliar modern ML or AI system** by identifying familiar components, explaining the purpose of its major design choices, and connecting it to the foundational ideas studied in the course.

Coursework and Assessment

Unless otherwise specified, assignments, projects, and project checkpoints will be due **Fridays 12:00 pm ET.**

Assignments [30%]: Periodic assignments involving coding exercises, written analysis, and brief oral check-ins. [Outcomes: a, c, d, e]
Project [25%]: A semester-long explanatory project, developed through several checkpoints and culminating in a final artifact and in-person presentation. [Outcomes: b, c, f]
Midterm [15%]: Week 8 examination (Oct. 23, 2026), covering pre-neural network material. [Outcomes: a, b, c]
Final Exam [25%]: Comprehensive examination covering theoretical foundations and practical applications. [Outcomes: a, b, c, f]
Participation [5%]: In-class activities, questions/discussion, and Canvas participation.

Late work policy: You have 48 total free late hours across the semester. After those are used, late work is penalized at 2% of the deliverable's grade per hour. If you anticipate needing more time for a particular assignment, please reach out before the deadline if possible.

Grading policy: Letter grades will be based on your overall percentage in the course (rounded to the nearest whole number, e.g. 89.5->90), using the ranges below. I may make small, class-wide adjustments at the end of the semester if an assessment proves unusually challenging or easy, but any changes will be applied consistently for all students.
A=90-100, B+=86-89, B=80-85, C+= 76-79, C= 70-75, D= 60-69, F: < 60.

Course topic schedule

Week	Topic
1	Overview, k-nearest neighbors, overfitting
2	Decision trees
3	Linear + logistic regression
4	Probability and maximum likelihood estimation (MLE); gradient descent
5	Regularization, hyperparameters, convergence; missed class (wellness day Oct 2)
6	Support vector machines, feature engineering, kernels; information theory
7	Generalization, expressivity, double descent
8	Midterm review; midterm
9	Backpropagation; multilayer perceptron
10	Convolutional, Recurrent neural networks
11	Transformer
12	Reading a real language model: nanoGPT
13	Unsupervised learning
14	Student presentations
15	Reinforcement learning overview; course synthesis

Course policies

Grade Corrections: Check the grades in course work and report errors promptly. Please try and resolve any issue within one week of the grade notification.

Incomplete: A grade of **I** (incomplete) is given in rare cases where work cannot be completed during the semester due to documented long-term illness or unexpected absence for other serious reasons. A student needs to be in good standing (i.e. passing the course before the absence) and receives a provisional I if there is no time to make up for the documented lost

time; an email with a timeline of what is needed to be done will be sent to the student. Note that an **I** must always be resolved by the end of the next semester.

Collaboration and External Resources for Assignments: Some homework will be challenging. You are advised to first try and solve all the problems on your own. For problems that persist you are welcome to talk to the course assistant or the instructor. You are also allowed to collaborate with your classmates and search for solutions online. But you should use such solutions only if you understand them completely (admitting that you don't understand something is way better than copying things you don't understand). Also make sure to give the appropriate credit and citation.

Accommodation of Disabilities: Office of Accessibility Resources and Services (OARS) offers long term and temporary accommodations for undergraduate, graduate and visiting students at NJIT. If you need accommodations due to a disability, please contact OARS via email at oars@njit.edu. The office is in Kupfrian Hall Room 201. For further information please visit the OARS office website at: https://www.njit.edu/accessibility/ Please note, if you are eligible for extra time and would like to use it in the final exam, please notify instructor and OARS at least two weeks prior to the exam so that accommodations can be made.

Student Absences for Religious Observations: NJIT is committed to supporting students observing religious holidays. Students must provide notification in writing of any conflicts between course requirements and religious observances, ideally by the end of the second week of classes and no later than two weeks before the anticipated absence. We will do our best to provide academically reasonable accommodations, allowing students to complete missed assignments, exams, quizzes, or other coursework within the term.

Generative AI Tools and Other External Resources:
You are welcome to use generative AI tools (e.g., ChatGPT) as part of your learning process unless instructed otherwise. You may not submit AI-generated work as your own. If generative AI contributes substantively to a submission, you must disclose the tool and briefly explain how it was used. You remain responsible for understanding and being able to explain everything you submit. Assignment-specific restrictions always take precedence, and generative AI may not be used during exams or oral assessments unless explicitly permitted. Failure to give appropriate credit when using the work of others (whether human or AI) is considered plagiarism, and may lead to disciplinary action under NJIT's Academic Integrity policy (see below).

Academic Integrity: Academic Integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: NJIT Academic Integrity Code.

Please note that it is my professional obligation and responsibility to report any academic misconduct to the Office of the Dean of Students. Any student found in violation of the code by cheating, plagiarizing, or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university. If you have any questions about the code of Academic Integrity, please contact the Office of the Dean of Students at dos@njit.edu.

Finally, your feedback is always welcome -- on lectures, assignments, or pacing. This course is designed to be iterative, and your input helps me improve it.