Introduction to Machine Learning
CS165B Fall 2023

Instructor: Prof. Yu-Xiang Wang

Lectures: Tuesday and Thursday 5:00 PM - 6:15 PM at CHEM 1171

Ed Discussion: https://edstem.org/us/courses/48177/discussion/
(We use Ed Discussion for Q&A, discussion, and most announcement!)

Gradescope: https://www.gradescope.com/courses/632349
(We will be using the gradescope for submitting code and project reports.)

Gauchospace: The Gauchospace is active, but it is only reserved for communications that are restricted to people who are officially registered.

1 Overview

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data," in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop a conventional algorithm for effectively performing the task (Wikipedia). We will cover the following topics:

This is an introductory course in artificial intelligence offered as part of a computer science curriculum, so we are not primarily concerned with the approaches of philosophy, cognitive science, etc., although we will come back to them from time to time. We will cover the following technical topics:

• Linear Algebra needed for machine learning
• Machine Learning Basic Concept
• Classification
• Regression
• Ensemble Methods
• Optimization
• Modeling principles
• Neural Networks, Deep Learning
Differences from CS165A: Note that CS165A also covers machine learning (ML). CS165B is different as it gives a more systematic introduction to machine learning. 165A takes a broader view and considers many other aspects of artificial intelligence (with or without learning). There isn’t a right order in taking these two courses. You can go either way. You will have more ideas on applications of ML for decision making if you have completed CS165A when taking CS165B; meanwhile you will gain deeper understanding in CS165A topics if you have taken CS165B.

2 What you will learn?

By the end of the course, you will understand what ML is all about and what it has contributed, and may contribute, to computing. You will have a working knowledge of the basic tools of machine learning, which are applicable to a wide range of computing problems. You will be able to solve problems using classification and clustering. You will have experience building models that, to some degree, can learn from experience.

This is not primarily a programming course - that is, the main goal is to learn the concepts, not to learn a language or particular programming techniques. However, coding examples of the concepts is the best way to demonstrate (and facilitate) your knowledge of them. Good programming practices (proper file structure, comments, etc.) are expected.

Quite different from a typical CS course, this is perhaps one of the first courses in your curriculum where you will learn how to consolidate what you have learned in mathematics, statistics and coding skills to solve problems. You will get some practices in simple mathematical derivations and translating the mathematical equations into codes in Python / Numpy.

3 Prerequisites

The formal prerequisite is Computer Science 130A (Data Structures and Algorithms I). This implies that you have studied topics such as algorithms, data structures, searching and sorting techniques, recursion, and induction, all of which are relevant to this course. Most importantly, you need to be able to think logically about problems and solution strategies, and you must be familiar enough with writing software to implement solutions on your own. If you have completed the CS pre-major courses (including the Math and PSTAT courses) and CS 130A then you will be very well-prepared.

4 Textbooks

There is no formal text book required, but we will assign readings from the following references books.

- “Machine Learning: The Art and Science of Algorithms that Make Sense of Data” by Peter Flach
- “Pattern Recognition and Machine Learning” by Chris Bishop
- “Dive Into Deep Learning”, by Zhang, Lipton, Li and Smola.

The latter two books are open access.

5 Assignments and Grades

The grades will be based on the following breakdowns.

- 30% Coding projects
• 30% Homework assignments
• 20% Midterm
• 20% Final Exam

Homework will include 4 assignments and 3 coding assignments. Unless otherwise instructed, all assignments must be submitted on the due date. No late submission is allowed — but only your best three homeworks will be included in your final evaluations.

You are encouraged to discuss homework assignments with classmates at a general level. However, you shall not share answers/code or collaborate on solutions unless otherwise directed to do so. All work turned in must be completely your own, including programming assignments. (See the Policy on Academic Integrity, below.)

There will be no makeup exams except for extraordinary circumstances.

6 Exams

The midterm and final will test your understanding of the materials from the lectures and discussion sections. They will be easier and shorter than the homeworks.

7 Logistics

Lectures and discussion section: The instruction will be in person as per UCSB guideline.

Office hours: Office hours are announced on the course website.

Attendance policy: The attendance to lectures and weekly discussion sections are required.

Late submission policy: No late days allowed.

8 Policy on Academic Integrity

Please read this section carefully.

The university, the department, and this instructor all take the issue of academic integrity very seriously. A university requires an atmosphere of mutual trust and respect. While collaboration is an integral part of many scholarly activities, it is not always appropriate in a course, and it is never appropriate unless due credit is given to all participants in the collaboration.

Here are some examples:

• Allowed: Discussion of lecture and textbook materials
• Allowed: Discussion of how to approach assignments, what techniques to consider, what textbook or lecture material is relevant
• Allowed: Collaboration on coding projects. You need to declare your collaborators, and describe what you get from the person who helped you, and each student still needs to write their own report / code independently.
• Allowed: Refer to online resources, but you need to cite the exact references.
• Not allowed: Turning in someone else’s work as your own, even with that person’s permission.
• Not allowed: Allowing someone else to turn in your work as his or her own.

• Not allowed: Turning in work without proper acknowledgment of the sources of the content (including ideas) contained within the work.

For some views on academic integrity at UCSB see the Academic Integrity page of the Office of Judicial Affairs.

Summary: Academic integrity is absolutely required - dishonesty (cheating, plagiarism, etc.) benefits no one and hurts everyone. Violations of these honor codes on academic integrity will be reported to the Office of Student Conduct. If you find yourself in such a situation, please contact the instructor. If you are not sure whether or not something is appropriate, please ask the instructor.

9 Code of conduct

The University of California, Santa Barbara has a general code of conduct for all students published here: http://www.sa.ucsb.edu/docs/default-source/student-conduct/conductofcode2017.pdf?sfvrsn=d3c07d4f_2

The computer science department’s commitment to Diversity, Equity and Inclusion is published here: https://cs.ucsb.edu/content/diversity-equity-and-inclusion

As a department, we holds students, staff, and faculty to the following standards:

• Treat all members of the academic community (students, staff, and faculty) with respect regardless of their experiences and background, including (but not limited to) their cultural backgrounds, socioeconomic status, disabilities, age, religion, sexual orientation, neuro(a)typicality, and gender identity.

• Physical or mental harm, sexual harassment, aggression, and derogatory language is not acceptable in any form.

• Respect the personal property of others and University resources. Unauthorized access, use, vandalism, or theft of equipment, computer servers, labs / offices / classrooms, etc. is prohibited.

• The exchange and challenge of ideas are done in a thoughtful, respectful and constructive manner.

• Disruption of departmental activities such as special events, talks, lectures, and meetings is not acceptable.

• Any violation of the given standards should be reported to the Computer Science chair and/or the CS diversity committee (diversity@cs.ucsb.edu). Consequences may include a formal warning, suspension, or expulsion from the University.

10 Students with Disabilities

If you are a student with a disability and would like to discuss special academic accommodations, please contact the instructor. In addition, students with temporary or permanent disabilities are referred to the Disabled Students Program (DSP) at UCSB. DSP will arrange for special services when appropriate (e.g., facilitation of access, note takers, readers, sign language interpreters). Please note that it is the student’s responsibility to communicate his or her special needs to the instructor, along with a letter of verification from DSP.
11 Copyright of course materials

My lectures and course materials, including presentations slides, written notes, recorded lectures, homework assignments and similar materials, are protected by U.S. copyright law and by University policy. I am the exclusive owner of the copyright in those materials I create. You may take notes and make copies of course materials for your own use. You may also share those materials with another student who is enrolled in or auditing this course. You may not reproduce, distribute or display (post/upload) lecture notes or recordings or course materials in any other way — whether or not a fee is charged — without my express prior written consent. You also may not allow others to do so.

If you do so, you may be subject to student conduct proceedings under the UC Santa Barbara Student Code of Conduct.