Course overview and logistics

CS861: Theoretical Foundations of Machine Learning

Kirthi Kandasamy
University of Wisconsin - Madison
Fall 2023

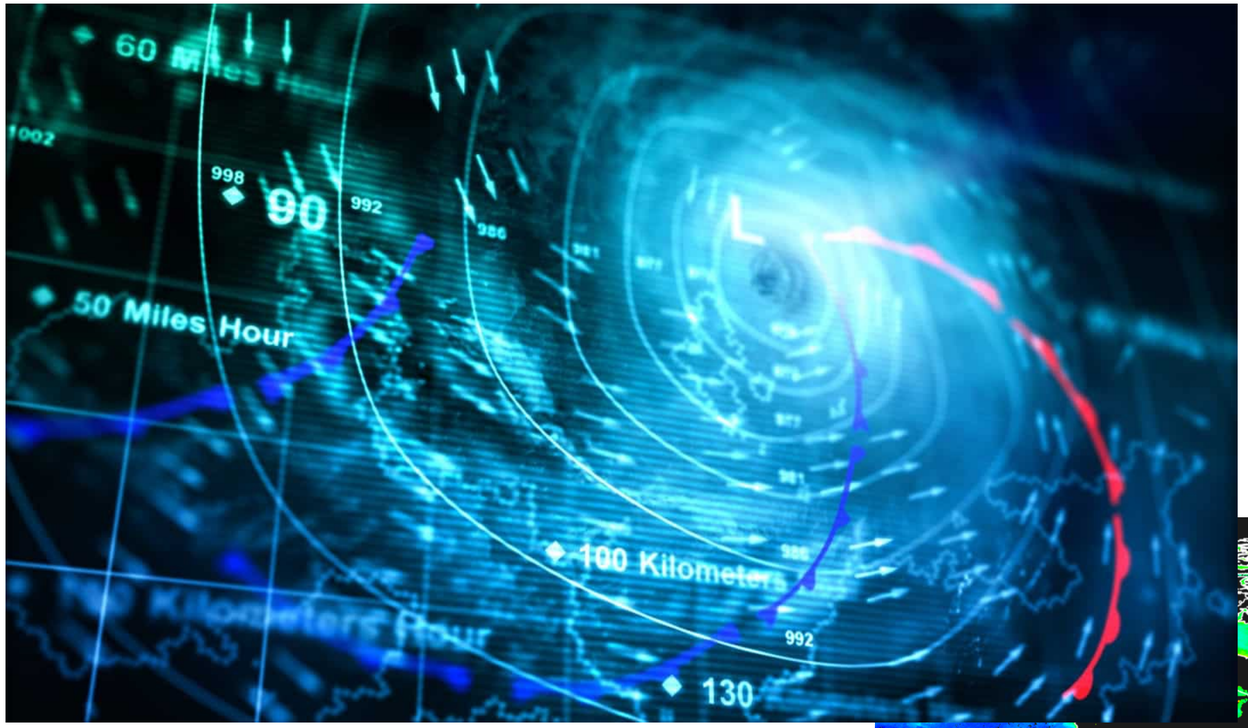
September 6, 2023

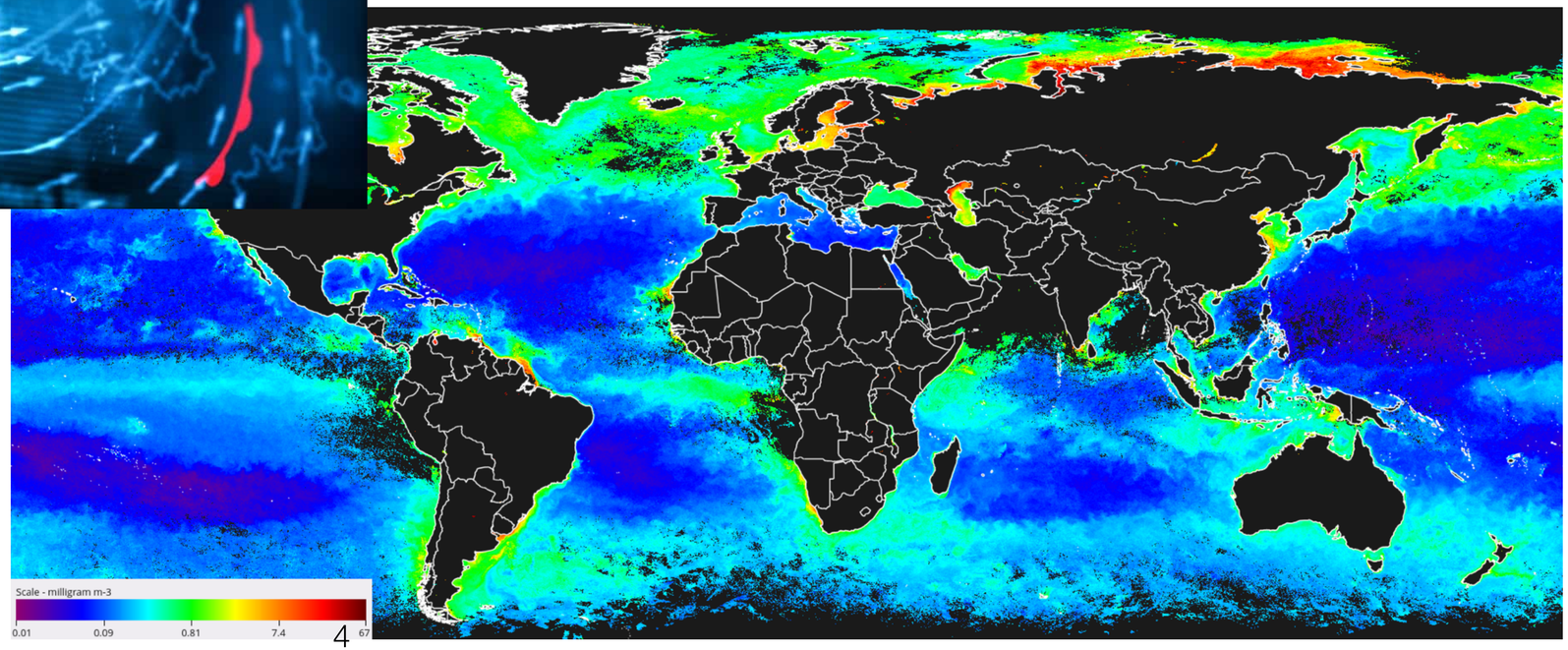
Machine learning is popular nowadays!

- *“A breakthrough in ML will be worth 10 Microsofts”*
  - Bill Gates
- *“ML is the new internet”*
  - Tony Tether, Director, DARPA
- *“AI will be the best or worst thing ever for humanity”*
  - Elon Musk

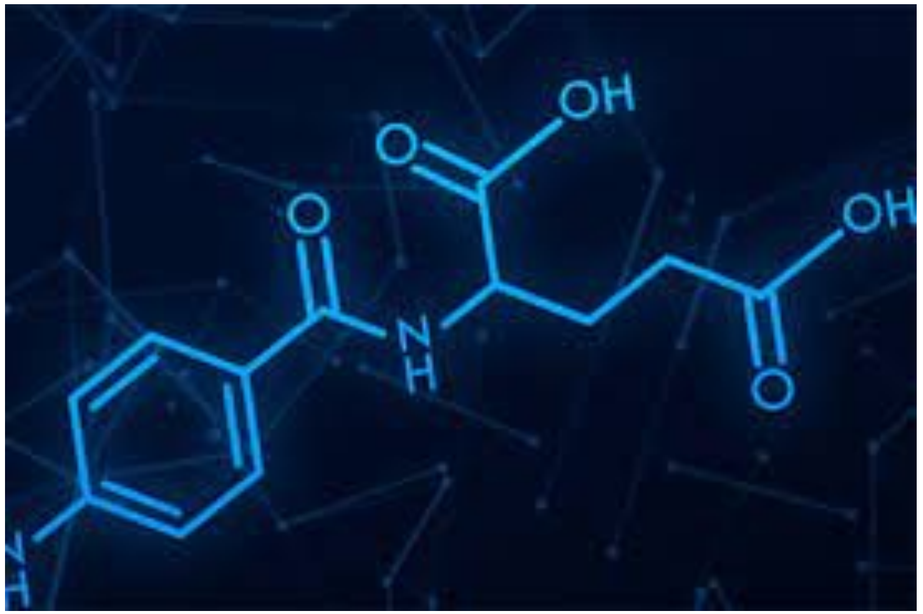
ML Application: Object detection & segmentation

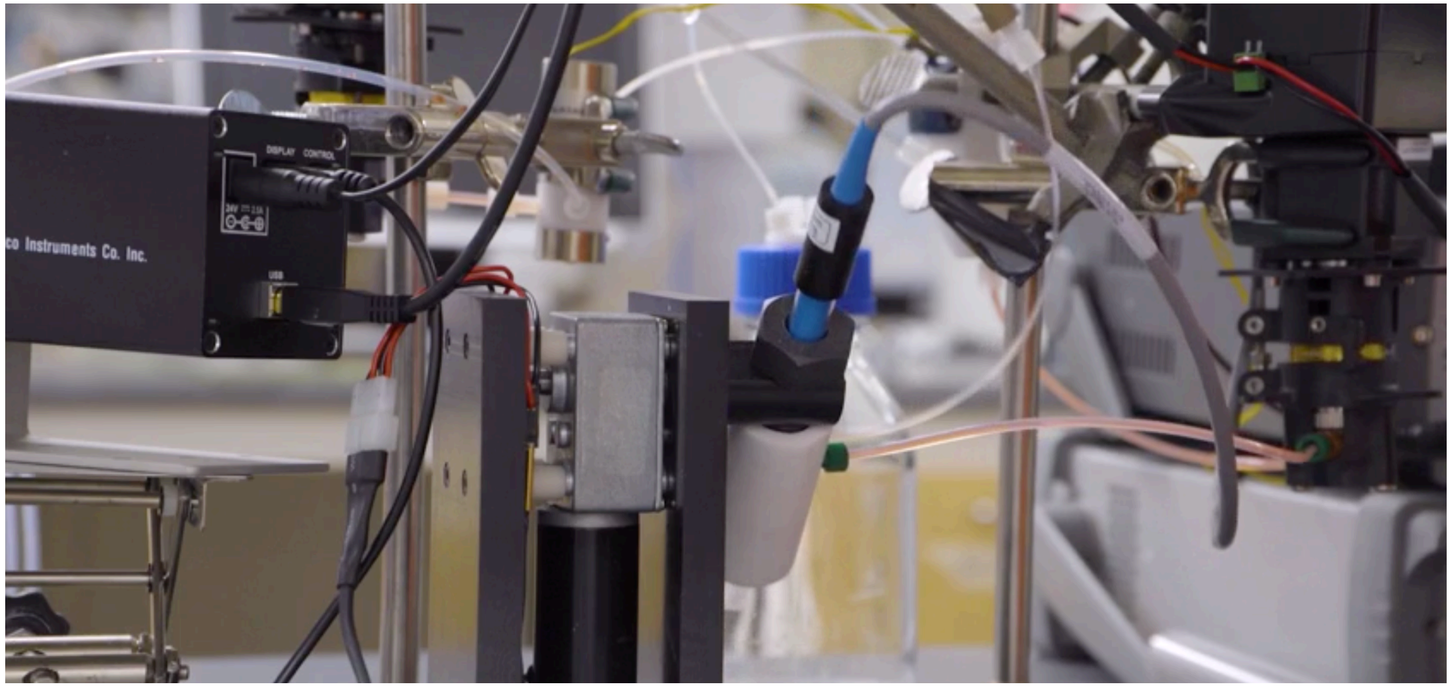
Weather forecasting & Climatology





Optimizing drugs and materials





Autonomous vehicles

Language generation (e.g GPT)

Q: What is your favorite animal?
A: My favorite animal is a dog.

Q: Why?
A: Because dogs are loyal and friendly.

Q: What are two reasons that a dog might be in a bad mood?
A: Two reasons that a dog might be in a bad mood are if it is hungry or if it is hot.

Q: How many eyes does a giraffe have?
A: A giraffe has two eyes.

Q: How many legs does a frog have?
A: A frog has four legs.

Q: Are there any animals with three legs?
A: No, there are no animals with three legs.

Q: Why don't animals have three legs?
A: Animals don't have three legs because they would fall over.

Image generation

Image to text generation

This class: *Theoretical Foundations* of ML

Why take this class? Why study ML theory?

1. Understand fundamental limitations about a learning problem.
    - Is it even possible to learn using data?
    - How much data do we need to learn?
    - What are the primary challenges when learning?

This class: *theoretical foundations* of ML

Why take this class? Why study ML theory?

2. Develop fundamental intuitions for designing learning algorithm
    - What is the “correct” approach to solve the primary challenges?
    - How do we trade-off between multiple challenges?
    - Will focus on simple (as opposed to *“realistic/practical”*) settings

3. It is fun!

Outline

1. Course logistics
2. Syllabus
3. Who should take this class? Prerequisites and expectations

Outline

1. **Course logistics**
2. Syllabus
3. Who should take this class? Prerequisites and expectations

Logistics: Lectures, OHs, Enrollment

- **Lectures**
  - MWF, 11-12.15am at Engineering Hall 3349
  - Will be on the whiteboard.
  - 27-30 lectures.

- **My office hours:** Wed 1.30 - 3pm at CS5375

- **Enrollment**
  - At capacity, but short waitlist.
  - Continue to come to class, some students will likely drop.

Logistics: Webpages

- **Course website**
  - https://pages.cs.wisc.edu/~kandasamy/courses/23fall-cs861
  - Information on logistics, syllabus, schedule, and grading

- **Piazza**
  - https://piazza.com/wisc/fall2023/csece861 (**access code: f23cs861**)
  - Ask public questions whenever possible.
  - Announcements, peer discussions on homework/lectures.

- **Canvas**
  - Homeworks, exams, and some announcements

Logistics: Scribing

- Each student will scribe ~2 lectures. Two students per lecture.
  - These details may change if enrollment drops.

- Sign up for scribing via the sign-up spreadsheet (see course website for link).

- **Instructions (see course website as well)**
  - Written in *full prose*, proof steps written in detail, intuitions explained well.
  - Prepare in Overleaf, and add me as a collaborator within 2 days
  - If you are unsure about taking the class or on the waitlist, sign up for after Oct 6.
  - If you decide to drop, *delete your name **and** email me*.

Logistics: Homework

- 4-5 Homeworks
- Physical copy due at the **beginning** of class (optionally, upload to canvas).

- Late submissions only for documented emergencies.

- 5 percent extra credit if you LaTeX your solutions.

- Homeworks will be *difficult*.
  - Expect to spend multiple hours/days on some problems.
  - Unless otherwise specified, you *are allowed* to collaborate with up to 2 classmates.

Logistics: Grades

- **Scribing: 10%**
- **Homeworks: 35%**
- **Exam: 30%**
  - Take-home exam, available from Tue 11/14 - Fri 11/17.
  - 48 hours to finish from start time.
- **Course project: 25%**
  - A final project. Should have a substantial theory-based component.
  - Project proposal due on 10/20. Final report due on 12/8.
  - I will reward high-risk projects.

Outline

1. Course logistics
2. **Syllabus**
3. Who should take this class? Prerequisites and expectations

Syllabus: Overview

1. PAC Learning
2. Statistical lower bounds
3. Online learning & bandits
4. Advanced topics

Syllabus: PAC Learning (4-6 lectures)

- Empirical risk minimization
- PAC Learning: realizable vs agnostic
- Radamacher complexity
- VC dimension

Syllabus: Statistical lower bounds (7-10 lectures)

- Average-risk optimality vs minimax optimality
- Minimax optimal estimators for point estimation
- From estimation to testing: Le Cam & Fano methods
- Applications
  - regression, classification, density estimation

Syllabus: Online learning (8-12 lectures)

- Learning from experts and the Hedge algorithm
- Adversarial bandits and the EXP-3 algorithm
- Stochastic bandits and the UCB algorithm
- Lower bounds for online learning and bandits

Syllabus: Advanced topics (~4 lectures)

- Learning in games
- Online learning and bandits in non-stationary environments
- Reinforcement learning

Outline

1. Course logistics
2. Syllabus
3. **Who should take this class? Prerequisites and expectations**

Target audience for the class

- Ph.D students doing research in theoretical **(statistical)** machine learning.

- **Background knowledge**
  - **Formal prerequisite:** CS761 or equivalent.
  - Strong background (intermediate-level graduate course) in calculus, statistics, and probability.

- **Who should not take this class**.
  - *“I want to learn about ML/AI”* (Take 540, 532)
  - *“I want to apply ML in an applied area of research*" (Take 760)
  - *“I want to learn take an introductory ML theory class”* (Take 761)

Homework 0

Three questions, going from easy to hard:

1. Mean estimation & concentration
2. Maximum risk
3. A simple bandit model and algorithm

Three Objectives

I. A preview of what's to come

II. Calibrate my teaching/expectation

III. Lets you assess if you are ready to take this class

Be good citizens!

1. Attend class, ask questions.
2. Take your scribing duties seriously
3. Respond to questions on Piazza.
4. Give me feedback about the course.
    - Are the homework problems useful?