COMS 4773 Fall 2026

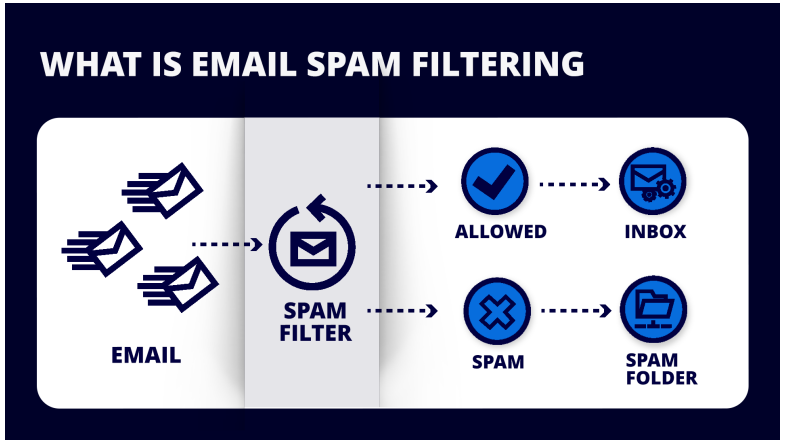
Daniel Hsu

What is machine learning (ML)?

- Examples:
  - Spam filtering (from email text)
  - Ad click prediction (from user profile and context)
  - Gene expression level prediction (from upstream DNA)
  - Best-next-move prediction (from state of chess board)
  - …
  - Programming-by-example
- Note:
  - This is not an introductory course in ML
  - We will not discuss application-oriented aspects of ML
  - See COMS 4771 and similar courses instead

Example: spam filtering

- General approach
  - Represent an email by several Boolean features
    - Does the email contain dollar signs?
    - Does the email have spelling errors?
    - Is the sender's address in your address book?
    - Is the sender's IP in a known botnet?
    - …
  - You label several emails as either "SPAM" or "NOT SPAM"
  - Provide these labeled emails to ML algorithm to derive a classification rule:
    - e.g., (has-dollar-signs AND (NOT sender-in-address-book)) OR sender-in-botnet ➔ SPAM
- How to derive the classification rule?
- Why should the rule work in the future?



What is learning theory?

- Design/analysis of ML algorithms/problems
  - Computational resources: running time, memory, ...
  - Information resources: sample size, rounds of interaction, ...
- Many different models for theoretical analysis
  - Online learning
  - Learning with queries
  - Statistical learning
  - Finding planted structures
  - …
- Each one tries to capture some essential aspects of "learning"
- Cross between theoretical computer science and statistics (and more)

Why study learning theory? (1/3)

- Relevance to machine learning practice
  [Breiman, 1995: "Reflections After Refereeing Papers for NIPS"]

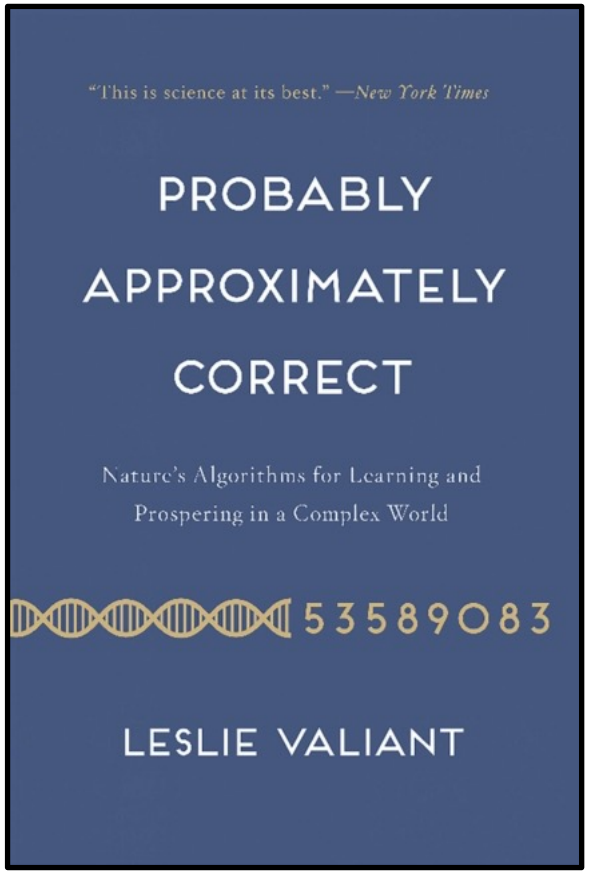
2. USES OF THEORY

- **Comfort:** We knew it worked, but it's nice to have a proof.
- **Insight:** Aha! So that's why it works.
- **Innovation:** At last, a mathematically proven idea that applies to data.
- **Suggestion:** Something like this might work with data.

Why study learning theory? (2/3)

- Insights into general phenomenon of learning [Valiant, 1984: "A Theory of the Learnable"]

ABSTRACT: *Humans appear to be able to learn new concepts without needing to be programmed explicitly in any conventional sense. In this paper we regard learning as the phenomenon of knowledge acquisition in the absence of explicit programming. We give a precise methodology for studying this phenomenon from a computational viewpoint.*



Why study learning theory? (3/3)

- Learning theory represents a very broad range of disciplines
  - Many different perspectives, broad range of topics of interest
  - Many techniques with wide applicability
- **Breiman's warning**:

> Mathematical theory is not critical to the development of machine learning.
>
> *But scientific inquiry is.*

> INQUIRY = sensible and intelligent efforts to understand what is going on.

- **Role of theorems/proofs/theory**:
  - can give evidence of understanding
  - can be useful for communicating ideas with clarity and precision

About this course

- **Website**: https://www.cs.columbia.edu/~djhsu/LT/
- **Waitlist**: managed by CS Advising (not me)
- **Topics**:
  - Online learning
  - Statistical learning
  - Other stuff (e.g., boosting, convex optimization, approximation theory, Gaussian concentration, spectral methods)
  - Significant overlap with COMS 4252, STAT 6203
- **Prerequisites**: mathematical maturity, probability, linear algebra, basic design/analysis of algorithms (see PS0)

Course requirements

- **Reading and lecture attendance**
  - No recordings; borrow notes from a classmate if you miss a lecture
- **Homework** (50% of overall grade)
  - Problem sets + peer evaluation
  - Graded only for completion, not correctness
- **In-class midterm exam** (25% of overall grade)
  - Closed-book, closed-notes, device-free, AI-free, etc.
  - Thursday, November 5
- **Final exam** (25% of overall grade)
  - Closed-book, closed-notes, device-free, AI-free, etc.
  - Tuesday, December 22 at noon (registrar's projection)

AI and collaboration policy

- **Using AI on problem sets**:
  - All problem sets in this course can be solved by AI
  - My gut feeling is that the educational value of the problem sets will be nullified if you use AI *at all* to solve them
  - Do not use AI for peer evaluation
- **Collaboration on problem sets**:
  - Okay to discuss problem sets with other students (in small groups)
  - But, strongly encouraged to try to do the problems yourself first
  - Must write up your solutions individually, without looking at anyone else's (partial or complete) write-up
- **Using AI on exams**:
  - Not permitted
  - Exams are closed-book, closed-notes, device-free, AI-free, etc.

Peer evaluation

- For some subset of the problem sets (and a subset of the problems), you will provide feedback on other students' write-ups
  - Practice evaluating mathematical arguments that may or may not be valid
  - We may throw in some purposely incorrect write-ups ☺
- Make sure your write-ups do not contain personally-identifying information (for anonymity)

First problem set

- First problem set: PS0
  - A lot of it should be familiar from prerequisite material
  - Some of it might be new, but not supposed to be hard to pick up
  - Use as self-assessment to decide whether this course is suitable for you
  - Submit on Gradescope