CS272: Foundations of Learning, Decisions, and Games

Lecture 1

Nika Haghtalab
UC Berkeley

Even your coffee is surprised that you are in class at 9am!

You do what you gotta do to learn in a 9am class!

Outline of Today

Who are we:

- Instructor: Nika Haghtalab
- Graduate Reader: Annie Ulichney

Annie Ulichney

Who are you?

What's this course about?

Some logistics

Gentle start to theory of machine learning

Learning and Learnability

One of the goals of theory of ML:

"What concepts can be learned from data, and with how many observations?"

An example of concept: "Familiar object, such as a table". [Valiant '84]

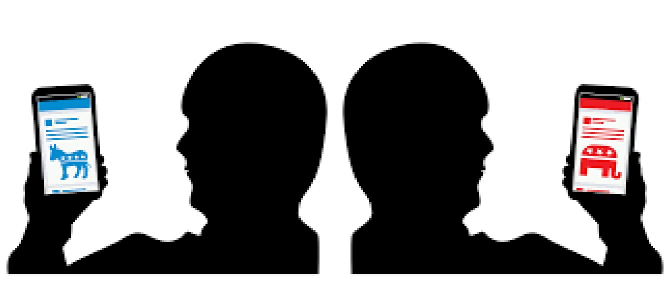
Formalized by: We have data $(x_i, y_i)s$ iid from distribution D and we want to learn the concept (function) table that is $f: X \to Y$.
Such that

$$\mathbb{E}[f(x) \neq y] \approx 0 \quad \text{with high probability}$$

Learnability for Today's World

Learning Algorithms

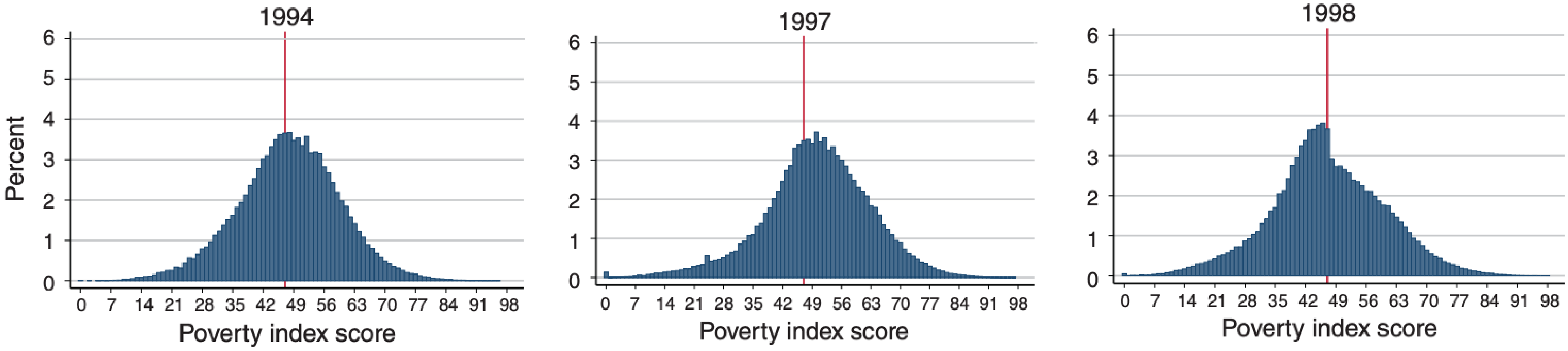
Environment



Case Study 1

In 1990s Columbian Gov. introduced a targeted social spending program

- Designed a “poverty index” meant to capture long-term conditions
- To identify “poor” population most in need of receiving assistance
- Using a census data
- Below the “poverty” threshold (red vertical line) = eligible for assistance



Case Study 1: What happened in late 1997?

The algorithm for "score" computation made public to municipal administrators and in pamphlets distributed to general public!

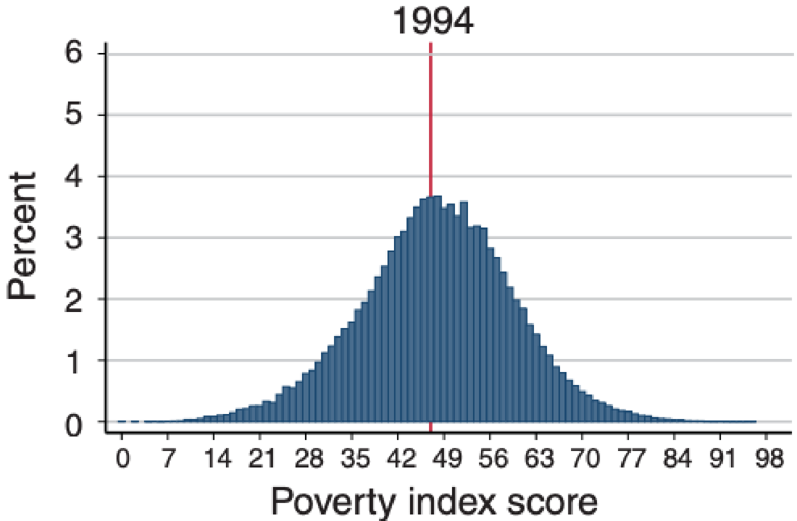
Data manipulations as many levels:

→ Respondent started to lie.

→ Surveyors manipulated data and interviews

→Corrupt politicians manipulated the data

Est. 7% of National Health and Social Security budget



Case Study 1: Takeaway

Distribution shift happen!

Sometimes as in social or strategic reaction to “learning” and “decision-making”

Learn more about this case study:

Manipulation of Social Program Eligibility†

By Adriana Camacho and Emily Conover*

We document how manipulation of a targeting system for social welfare programs evolves over time. First, there was strategic behavior of some local politicians in the timing of the household interviews around local elections. Then, there was corrupt behavior with the sudden emergence of a sharp discontinuity in the score density, exactly at the eligibility threshold, which coincided with the release of the score algorithm to local officials. The discontinuity at the threshold is larger where mayoral elections are more competitive. While cultural forces are surely relevant for corruption, our results also highlight the importance of information and incentives. (*JEL* D72, I32, I38, O15, O17).

Case Study 2

In 2018 Zillow expanded from publishing price predictions ("Zestimates") into directly buying and reselling homes.

- Zillow used its Zestimates to make *real financial decisions*.
- Their strategy: buy homes near Zestimate, light repairs, and resell for profit.

Zestimates were really accurate:

- **average prediction error** of about 1% on houses sold and 4% on all of the market.
- Their smaller scale deployment did very well in 2018-2020 producing twice more profit than they had expected

But .... by 2021, Zillow lost 500M dollars and had to fire 25% of workforce just to survive!

Case Study 2: What went wrong with Zestimates?

A lot of issues contributed to this:

1. There was unforeseen market shift caused by the pandemic (2020–2021) causing significant **distribution shift**.
2. More fundamentally: **Zestimate wasn't calibrated**

- Average error low, but systematically overestimating in hot markets
- Owners know more about their house (like insider traders!), so they sold when Zestimate was an overestimate, but not when it was underestimate.

Why didn't they catch this earlier?

- Priced were increasing rapidly in 2018-2020, so even overestimates didn't make Zillow lose money. Cooling market surfaced the issues.

Case Study 2: Takeaway

1. Average accuracy isn't enough, Calibrated predictions needed for decisions to be good.

2. Historical success doesn't mean future success! Unless, we understand the contributing factors to the outcomes.

Commentary: How homeowners defeated Zillow's AI, which led to Zillow Offers' demise

BY OREN ETZIONI on January 31, 2022 at 7:00 am

THE WALL STREET JOURNAL.

Single-family homes in Del Mar, Calif. BING GUAN/BLOOMBERG NEWS

By *Will Parker* Follow and *Konrad Putzier* Follow

Nov. 17, 2021 11:29 am ET

Case Study 3

Healthcare systems use algorithms for population health management:

- Identify patients with higher needs and give them more resources.
- They use **past data** to build a **risk score** of future health care needs
- Past data = Cost of care received in the past
- Ensure that the risk score is **calibrated** for the population

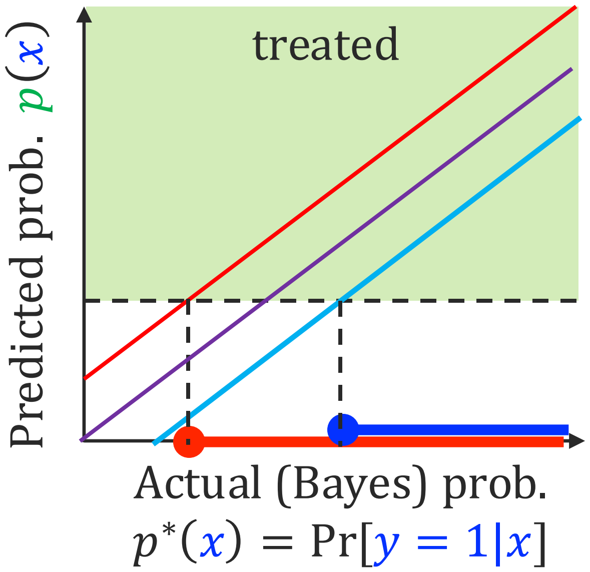
At a given risk score, Black patients are considerably sicker than White patients.

Case 3: What's going wrong?

The risk score is calibrated on the population as a whole.
On average the "purple" line, is a good score.

But we can still have mis-calibrated across sub-populations.
The red population and blue population treated differently

Lack of predictions across two groups means that predictions systematically mean something different about the two groups.



Why is this happening?

- Past health care cost is a bad proxy for current health conditions,
- Because of historical unequal access to care mean.

Case Study 3: Takeaway

Calibration might not be enough!

We might need more fine-grained performance guarantees.

How fine-grained? Depends on the downstream decisions and goals

Learn more about this study:

RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2]*, Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5]*†

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

Case Study 4

Markets and competition are great for the customers:

- If one firm raises prices, customers move to the other and punish the higher seller
- Prices fall toward true cost and benefit consumers

Competition is an important part of economy.

But increasingly market decisions are made by algorithms, not people

- Firms use **ML/RL-powered pricing** that automatically adjusts prices in real time.
- Trading firms use AI-powered strategies and bots
- Landlord use platforms to set their rents

Do algorithms keep markets competitive—or can they learn to collude?

Case Study 4: What happened?

Classical "collusion" happens when firms communicate and make deals.

Algorithms that learn from the same data, implicitly communicate their strategies.

Researchers tested simple **RL pricing agents** (Calvano et al., *AER* 2020).

- Instead of competing prices down, the algorithms learned to keep prices high.

→When one reduced price, others punished with a price war

→ Then they all gradually returned to high, collusive prices.

→Sometimes, because of the particular exploration strategy of the online algorithms

Case Study 4: Takeaway

Dynamics between learning algorithms reveal different equilibria.

- Historically, Nash equilibria were proxy of market behavior
- Now, need to understand algorithmic dynamics and equilibria.

PRESS RELEASE

Justice Department Sues RealPage for Algorithmic Pricing Scheme that Harms Millions of American Renters

Friday, August 23, 2024

Artificial Intelligence, Algorithmic Pricing, and Collusion†

By EMILIO CALVANO, GIACOMO CALZOLARI, VINCENZO DENICOLÒ, AND SERGIO PASTORELLO*

Increasingly, algorithms are supplanting human decision-makers in pricing goods and services. To analyze the possible consequences, we study experimentally the behavior of algorithms powered by Artificial Intelligence (Q-learning) in a workhorse oligopoly model of repeated price competition. We find that the algorithms consistently learn to charge supracompetitive prices, without communicating with one another. The high prices are sustained by collusive strategies with a finite phase of punishment followed by a gradual return to cooperation. This finding is robust to asymmetries in cost or demand, changes in the number of players, and various forms of uncertainty. (*JEL* D21, D43, D83, L12, L13)

Foundations of Learning, Decisions, and Games

Online Learning
Dealing with unpredictable or adversarial distribution shifts

Dynamics between learning Algorithms and Equilibria

Statistical Learning

Game Theory and Equilibria

Stackelberg Games
Preparing for strategic responses

Multi-distribution and Muti-objective learning
Preparing for adversarial shift

Trustworthy Predictions
for downstream decision-making

Calibration and Scoring Rules

Fine-Grained guarantees
Multi-calibration and omni-prediction

Logistics

Lecture Delivery

We will use slides occasionally.

But, most of the lectures will be on the board.

→I will release lecture notes for the first few lectures.

→The rest of the lectures will be scribed by you!

Sign up sheet will be on Ed, please sign up by 09/01.

What you need for success in the course

Interest and hard work!

This course:

- Theory! Motivated by needs of the real world applications.
- Understanding the mathematical foundations.
- No programming skills needed.å
- You need to be comfortable with understanding/writing mathematical proofs and abstractions.

Having taken courses in 2 of

1. Stats and Prob., e.g., STAT205/210
2. Economics, e.g., ECON207
3. Algorithms, e.g., CS270
4. Optimization, e.g., EE 227
5. Control theory, e.g., EE 221A

Course Material

No required textbook.

→ We will rely on in class presentations (white board/slides), lecture notes, scribe notes, and other online material.

Great additional resources (see the website for links):

- Shalev-Shwartz, Ben-David, "Understanding Machine Learning - From Theory to Algorithms", Cambridge University Press.
- Blum, Hopcroft, Kannan, “Foundations of Data Science”, Cambridge University Press.
- Bousquet, Boucheron, Lugosi, “Introduction to Statistical Learning Theory”, Springers.

Course and Grade Components

- Deliverables
  - Homeworks (0-5) (42% of Grade)
  - Final Project and milestones (40% of Grade)
  - Project peer review (8% of Grade)
  - Scribing (7% of Grade)
  - Class participation (3% of Grade)
  - No final exam

Check out the homepage for detailed policy.

Homework Assignments

Assignments

- 6 written homeworks, total 42% of total grade.
- Tentative due dates already online.
- Include algorithm design, analysis, proofs, etc. No programming.

Policies

- Can be done in teams of up to 2.
- Everybody has 5 “free” late days. When you run out of late days, you’ll loose 1% per day from the total homework grade. No late days for HW0.
- No assignments will be accepted after the solutions have been made available (typically 2-5 days after deadline).
- Do not post the homework or your solutions publicly.

Final Project

The class has a final project:

- A project proposal Version 1 due mid way through the semester
- Peer Review of project proposal
- Project proposal Version 2!
- You will have a final report and likely a poster session.

Scope of project:

- Original research artifact, with at least some theorems and proofs.
- Suitable for workshop or conference submission.
- Purely survey papers are not in scope.

Projects can be done in groups of 1-3. I recommend teams of 2! Reserve 3 for very substantial scope.

Scribing

Full policy will be announced after the size of the class is finalized (after add/drop).

Scribing:

- Each student will be responsible for scribing 1-2 lectures.
- A lecture can be scribed by 1-2 students.
  - →Two students can scribe a lecture together only if all lectures already have a volunteer

Policy:

- We need fast turnaround!
- Within 2 weekdays, prepare a draft and share it with the instructor and TA. You have two more weekdays to prepare the final version.
- We aim to post the scribed notes within a week of the lecture.

How to Get in Touch

Course Homepage (slides, notes, additional resources, course policies, etc.)

- https://people.eecs.berkeley.edu/~nika/courses/cs272/f25

Ed for forum, communication with instructor and staff.

Gradescope for homework submission

Office Hours

- Nika: will be announced
- Annie Ulichney (by appointment, only to discuss grading)

Announcements

HW 0 will be posted next week and due a week later.

→Designed to testing your preparedness and not directly based on the material I'll be teaching.

→If you struggle with it, it doesn't mean you can't take the course. Just that you may have to work harder or refresh your knowledge.

Next two lectures

→I'm traveling.

→Lectures will be taught remotely or in–person by substitute instructor.

→Check Ed for more information.