Causality and Machine Learning
Lecture 1 – Introduction

Prof. Makar

Today's agenda

- Course overview and goals
- Course logistics
- Causation vs. association
- The causal inference pipeline
- Review of pre-requisites

Course overview

- Causality and prediction using ML
- What is causal inference
- Advantages/disadvantages of using ML methods for causal inference
- Addressing limitations of ML-based predictive models using ideas from causality

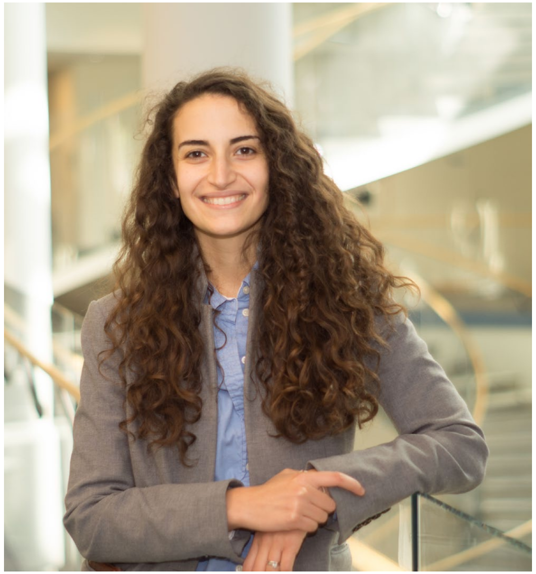
Course goals

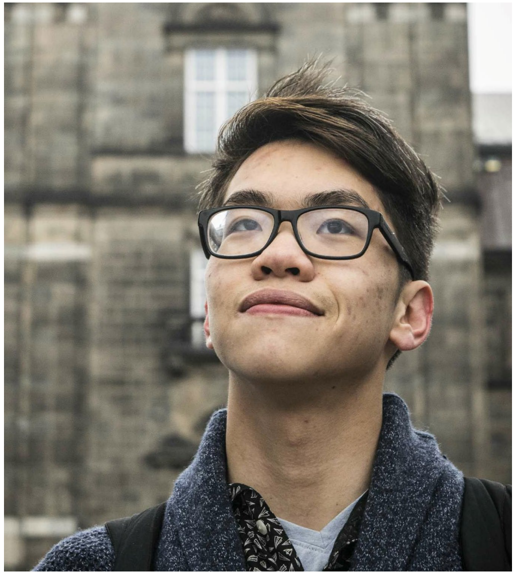
- Construct and implement causal models
- Reason about the reliability of these models with some rigor
- Comfort and familiarity with rigorous proofs
- Developing better predictive models by incorporating causal reasoning
- [Informal] Community building

Poll

- Who here is:
  - From CS? ECE? Stats? SI? Econ
  - A PhD student? Master student?
  - Has taken a causality class before?
  - Knows what a potential outcome is?
  - Knows what d-separation is?

Course logistics: class staff





Instructor: Professor Maggie Makar
Email: mmakar@umich.edu
OH: Wednesdays 1:40-2:40
Sign up required (link in syllabus)

GSI: Trenton Chang
Email: ctrenton@umich.edu
OH: Monday, ~~2:30 – 3:30~~ 3–4
Sign up required (link in syllabus)

Class e-mail: eecs598-cml-staff@umich.edu

Class logistics: Enrolling in the class

- Get on the waitlist
- Complete this form
- You will get temporary canvas access

http://bit.ly/eecs598waitlist

Class logistics: material

- Canvas
  - Class recordings
  - HW/project prompts
  - Announcements
- HW through gradescope (through canvas)
- Piazza for discussions (through canvas)

Grading

- Class participation 10%
- HW 39% 4
- Project 50%
  - Broken down into different components
- Course evaluations 1% (0.5% each)

- Late days:
  - 4 total
  - 3 max per one HW
  - 2 max for project final draft
  - 1 max for project draft 1
  - 0 max for project presentations

Class logistics: tentative schedule

Date	Topic	Released	Due @ 10pm
Mon Aug 28	Lec 1: Introduction	HW0	
Wed Aug 30	Lec 2: Causal notation and language		
Mon Sep 4	LABOR DAY		
Wed Sep 6	Lec 3: ATE under randomization	HW1	HW0
Mon Sep 11	Lec 4: ATE with covariates -- plug-in estimators		
Wed Sep 13	Lec 5: ATE with covariates -- doubly robust estimators		HW1
Mon Sep 18	Lec 6: ATE under measured confounding		
Wed Sep 20	Lec 7: ATE under measured confounding -- doubly robust estimators	HW2	
Mon Sep 25	Lec 8: CATE under measured confounding		
Wed Sep 27	Lec 9: CATE under measured confounding		HW2
Mon Oct 2	Lec 10: Causally motivated predictive models -- robustness		
Wed Oct 4	Lec 11: CMPM -- robustness	Project prompt released	
Mon Oct 9	Lec 12: CMPM -- robustness		
Wed Oct 11	Lec 13: CMPM -- efficiency		
Mon Oct 16	FALL BREAK		
Wed Oct 18	Lec 14: CMPM -- fairness		Proposal due
Mon Oct 23	Lec 15: CMPM -- fairness		
Wed Oct 25	Project one-on-ones no lecture		
Mon Oct 30	Lec 16: ATE estimation with hidden confounding		
Wed Nov 1	Lec 17: CATE estimation with hidden confounding		Draft 1 due
Mon Nov 6	Lec 18: CATE estimation with hidden confounding		
Wed Nov 8	Lec 19: Proximal causal inference	HW3	Draft 1 reviews due
Mon Nov 13	Lec 20: Instrumental variables		
Wed Nov 15	Lec 21: Mediation		HW3
Mon Nov 20	Lec 22: Causal discovery		
Wed Nov 22	THANKSGIVING BREAK		
Mon Nov 27	Lec 23: Causal discovery		Project final drafts due
Wed Nov 29	Project presentations		
Mon Dec 4	Project presentations		

Causal

Pred

Causal

Miscellaneous

Van der vaart 2000
Boos & Stefanski

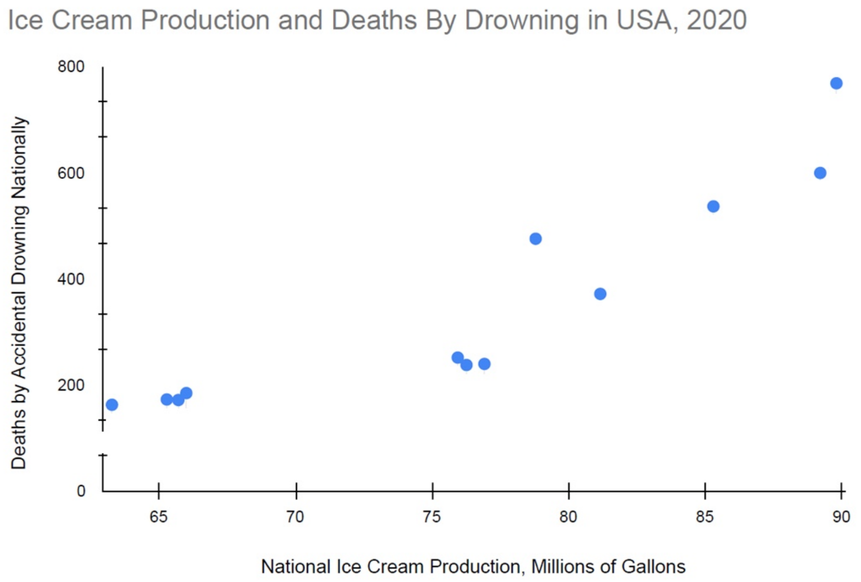
- Brush up on stats/probability
- This class: heavier on theory lighter on coding
- No single textbook
- Use of generative AI

To-do's this week

- HW0 was released on Canvas, due Wednesday Sept ~~10~~ 6 @ 10pm
- We want to know you!
  - Send an email to eecs598-cml-staff@umich.edu with your name phonetically spelled out and a recent picture.
  - Include [student-intro] in the subject line
- Read through the syllabus

Causation vs association

- Who is more likely to shoot 3 pointers?
- Does drip spray increase chance of shooting 3 pointers?



- $\hat{Y} = \underline{\hat{\beta}_1} X_1 + \underline{\hat{\beta}_0} X_0$

Why causal inference?

- Does including the citizenship question on the US census (ACS) lead to a suppression in answers from non-citizens?

Why causal inference?

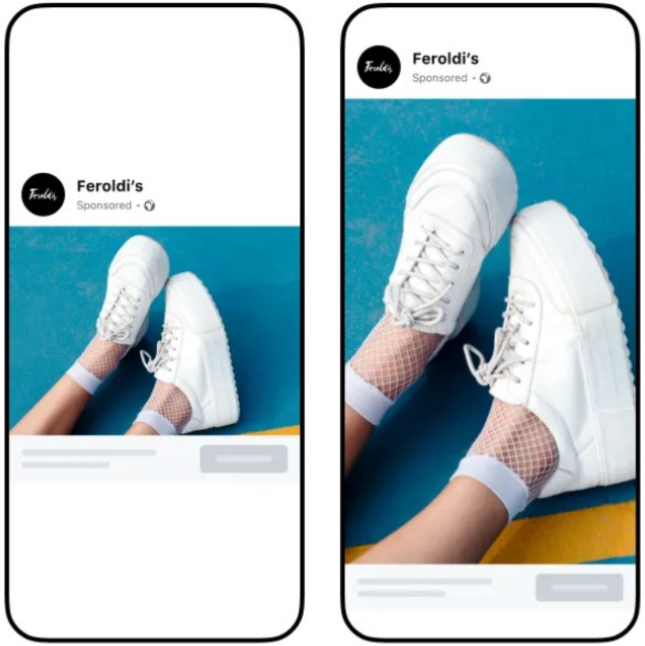
- Does OxyContin lead to an increase in mortality and substance-abuse?

Gruber reached the following principal conclusions: First, "[t]here is a direct, causal relationship between defendants' shipments of prescription opioids and the misuse and mortality from prescription opioids, with geographic areas that received higher volumes of per capita shipments of prescription opioids experiencing significantly higher rates of opioid related misuse and mortality, including the Bellwether jurisdictions." *Id.* at ¶ 16. Second, "[t]here is a direct, causal relationship between defendants' shipments of prescription opioids and the misuse of and mortality from illicit opioids, including heroin and fentanyl, which accelerated rapidly after 2010." *Id.* Third, "[t]he significant increases in all-opioid mortality (i.e., mortality from both prescription and illicit opioids) are largely unrelated to trends in non-opioid drug overdoses, changes in population demographics, or local economic conditions." *Id.*

ORDER DENYING MOTION TO EXCLUDE GRUBER

Why causal inference?

- Does a horizontal view of my product lead to better social media engagement?



Horizontal video

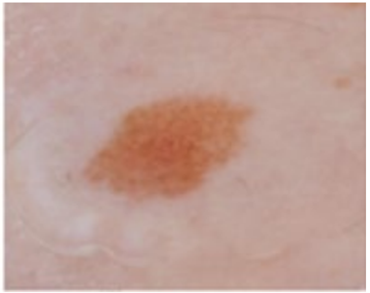
Vertical video

Why causally-motivated predictive models?

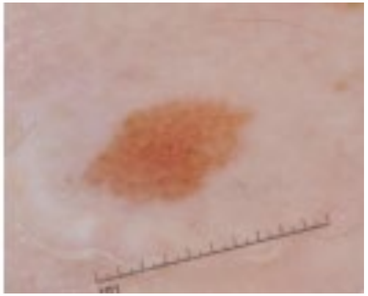
- Robustness to shortcut learning/distribution shifts
- Generalizing to unseen distributions
- Finite sample efficiency

$P(\text{Cow}) = 0.99$

$P(\text{Cow}) = 0.01$



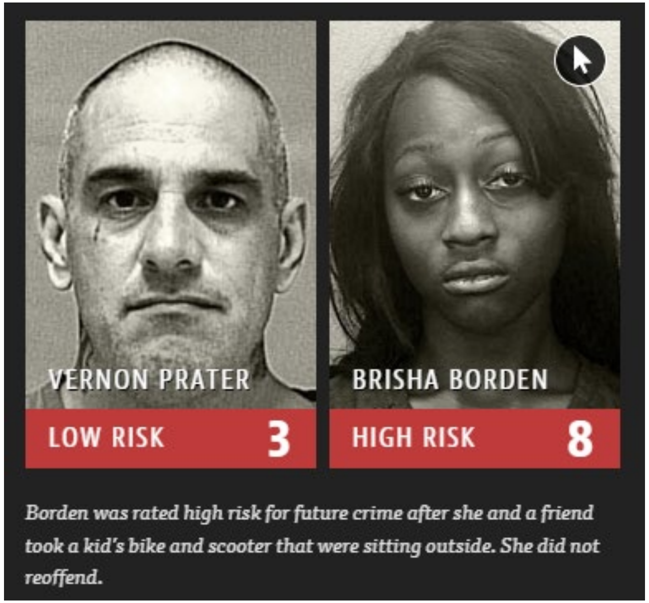
$\mathrm{P}(\text{Malignant}) = 0.01$



$\mathrm{P}(\text{Malignant}) = 0.99$

Why causally-motivated predictive models?

- Enforcing notions of fairness



Diff bet causal inf & predictive model:

(I) (1) PM → reason about the world as it is

(2) CI → reasons about the world as it would be under an intervention.

(II) Diff in "CI" pipeline

(III) Diff in notation & language.

Causal pipeline

(1) Formalize the target causal estimand

- Would an NBA player with 5 yrs of experience be able to shoot 3pointers better if given dry spray
- what's the avg # of drownings if we ban ice cream production?

(2) Establishing identifiability [or lack thereof]

- Expressing the target estimand in terms of the observed data
- OR reaching conclusion that estimand can't be expressed in terms of obs data.
- To establish identifiability:

(1) Codify our causal knowledge (assumptions)

The month of the year (the temp) is a common cause of drowning & ice-cream consumption.

(2) Ask if the obs data satisfies / enables identification.

(III) Estimation.
Using statistical tools (ML).

Review of stats / Prob

Ex 1: $X_i \sim N(\alpha, 1.2)$

Ex 2: $X_i \sim \mathbb{P}$

Ex 3: $Y_i = \beta_0 + \beta_x X_i + \epsilon_i \qquad \epsilon_i \sim N(0,1)$

Estimand: Property of the population distribution that is the goal of our inference.

Ex 1. α

Ex 2: $\mathbb{E}_{\mathbb{P}}[X_i]$

Ex 3: $\beta = [\beta_0, \beta_x]$.

Estimator: RV capture our guess about the value of the estimand.

Ex 1: $\hat{\alpha} = \frac{1}{n} \sum_i X_i$

Ex 2: $\hat{\mu} = \frac{1}{n} \sum_i X_i = \mathbb{P}_n(X)$

Ex 3: $\hat{\beta} = \underset{\beta \in \mathbb{R}^2}{\operatorname{argmin}} \frac{1}{n} \sum_i (y_i - \beta_0 - \beta_x x_i)^2$