Machine learning, causal inference, and economics

Maximilian Kasy

Department of Economics, University of Oxford

Winter 2026

Introduction

- What is the intersection of machine learning, causal inference, and economics?
- One possibility:
  - Supervised learning for first-stage estimators
  - in double-robust procedures,
  - under conditional exogeneity.
- But there is a lot more out there!
- This talk:
  1. 6 arguments about causality, ML, and econ.
  2. Elaborated in the context of binary choice.
- Goal: **New research agendas**.

6 Arguments

Formalizing these arguments

Argument I

- Common view:
  - Machine learning (ML) is purely correlational.
  - Causality needs to be introduced as a new concept.
- However:
  - Causality is already part of the core of ML.
  - Reinforcement learning (RL), including bandits.

Argument II

- Causal inference remains implicit in RL.
- Causal effects
  - of algorithmically chosen actions
  - on rewards.
- The chosen actions are by construction exogenous,
- conditional on the available information.

Argument III

- Notions of causality in econometrics, biostatistics, computer science:
  - Structural functions,
  - potential outcomes,
  - do-calculus.
- Each of these notions: Based on exogenous intervention.
- $\Longrightarrow$ Causality and decision-making are intimately connected.
- In RL, we can identify exactly the causal effects needed.
  - What you need is what you get!

Argument IV

- In RL, the focus is on the effect of actions on *observable rewards*.
- In welfare economics, *utilities* are *not observable*.
- But we can infer utility from observed behavior!
- E.g. Consumer surplus $=$ integrated demand.
- This leads to harder exploration / exploitation tradeoffs, relative to bandits.

Argument V

- The "fundamental problem of causal inference:"
  - We only observe one of the potential outcomes,
  - corresponding to the realized intervention.
  - Counterfactuals are "missing data."
- But:
  - In economic settings, causal inference does *not* always require intervention.
  - Mechanism design: Preference elicitation to recover response functions.

Argument VI

- Peoples' welfare also depends on the actions of others.
- Externalities are preferences over such actions of others.
- Behavior only reveals preferences over one's own actions.
- Eliciting externalities is still possible, but requires more elaborate mechanisms.

6 Arguments

Formalizing these arguments

Example: Binary choice

- Individuals i.
- Price $X_i \in [0, 1]$.
- Willingness to pay V_i.
- Binary choice $Y_i \in \{0, 1\}$, where

$$Y_i = \mathbf{1}(X_i \leq V_i).$$

- Potential outcomes $Y_i(x) = \mathbf{1}(x \leq V_i)$.
- Covariates W_i.

Causal inference in the binary choice model

- Exogenously (randomly) assign prices X_i, independently of V_i.
- Identify average demand (average structural function) via

$$E[Y(x)] = E[\mathbf{1}(x \le V)] = E[\mathbf{1}(x \le V)|X = x] = E[Y|X = x].$$

Causality versus machine learning?

Common framing:

- Canonical ML does not consider causality.
- (Supervised) learning is about prediction: Estimating

$$g(x, w) = E[Y|X = x, W = w],$$

or

$$p(x, w) = P[X = x|W = w].$$

- Causal inference is about identifying treatment effects,

$$ATE = E[Y(x^1) - Y(x^0)].$$

Supervised learning for first stage regression

- Under conditional independence, $V \perp X|W$, supervised learning can be used for first stage regressions and propensity scores.
- These can then be used to estimate treatment effects.
- E.g. via plug-in estimation,

$$ATE = E[g(x^1, W) - g(x^0, W)].$$

- Alternatively, using double-robust estimating equation,

$$ATE = E\left[g(x^1, W) - g(x^0, W) + \frac{\mathbf{1}(X = x^1)(Y - g(x^1, W))}{p(x^1, W)} - \frac{\mathbf{1}(X = x^0)(Y - g(x^0, W))}{p(x^0, W)}\right].$$

Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey, and Robins [2018].

Bandits and reinforcement learning

- Supervised learning, un/self-supervised learning:
  - Data (W, X, Y) are observed independently of algorithm choices.
- Active learning:
  - Data observability depends on algorithm choices.
  - Multi-armed / contextual bandits, reinforcement learning.
- For active learning, algorithms iteratively
  - choose action X_i,
  - then observe the reward R_i for the chosen action.
- Experimentation and estimation.

Example: Monopoly pricing

- Profits $R_i = Y_i \cdot X_i = \mathbf{1}(X_i \leq V_i) \cdot X_i$.
- Potential profits: $R_i(x) = \mathbf{1}(x \leq V_i) \cdot x$
- Contextual bandit: Iteratively
  (1) observe W_i, (2) choose X_i, (3) observe R_i.
- Goal: Maximize $\sum_i R_i$.
- Tradeoff:
  - Exploration: Experiment to learn the mapping $(x, w) \to E[R(x)|W = w]$.
    This is causal inference!
  - Exploitation: Choose X_i to maximize $E[R_i(x)|W_i]$.

Kleinberg and Leighton [2003].

Interventions, experiments, and causal effects

- The experimental ideal (physics):
  - Prepare a system, observe what happens.
  - Change something about the preparation, observe how outcomes change.
  - *Exogenous intervention* thus reveals counterfactual outcomes.
- Life sciences, social sciences:
  - Unobserved *heterogeneity* is pervasive.
  - $\implies$ Controlled preparation is impossible.
  - But we can repeatedly intervene and observe average outcomes.
  - Exogenous intervention reveals *average* counterfactual outcomes.

Experiments and decision problems

- We define causality based on this reference point of exogenous intervention!
- Similarly, decision problems consider
  - the expected reward (outcome)
  - when exogenously choosing an action (intervention).
- Experimental causal inference and decision-making are thus closely related.

Observed profits versus unobserved welfare

- Reward $R_i(x)$ in RL is observed when choosing x:
  - Winning a game of go or chess.
  - Patient mortality.
  - Profits for price x.
  - Ad clicks.
- (Social) welfare in economics is unobserved:
  - $SWF = \sum_i \omega_i U_i$.
  - Welfare weights ω_i.
  - Utility U_i of individual i.
- But utility U_i can indirectly be inferred from observed behavior.

Utility and behavior in the binary choice model

- Recall that $Y_i = \mathbf{1}(X_i \leq V_i)$.
- Individual Utility $U_i = \max(V_i - X_i, 0)$.
- We can rewrite this as an integral,

$$U_i = \int_{X_i}^{1} \mathbf{1}(x \leq V_i)dx = \int_{X_i}^{1} Y_i(x)dx.$$

- For randomly assigned $X_i \perp V_i$:

$$E[U_i(x)] = \int_{X_i}^{1} E[Y_i | X_i = x']dx'.$$

- Consumer surplus is area under the demand curve.

Example: Optimal taxation

- Social welfare and potential social welfare:

$$S_i = \underbrace{Y_i \cdot X_i}_{\text{Tax revenue}} + \lambda \cdot \underbrace{\max(V_i - X_i, 0)}_{\text{Individual utility}}.$$

$$S_i(x) = Y_i(x) \cdot x + \lambda \cdot \int_x^1 Y_i(x')dx'.$$

- To know welfare difference between two tax rates x_1, x_2, need to know function $E[Y_i(x)]$ for all $x \in [x_1, x_2]$.
- This changes exploration / exploitation tradeoff:
  - For bandits, only need to explore potentially optimal actions.
  - For social welfare, might need to explore suboptimal actions.

Cesa-Bianchi, Colomboni, and Kasy [2025].

Mechanism design: Causality without intervention

- Above I claimed that causal identification requires intervention (exogenous variation).
- Is that true? Consider the following mechanism:
  - Commit secretly to a price X_i.
  - Ask individual to report their willingness to pay V_i.
  - If $X_i \leq V_i$, implement $Y_i = 1$, and charge X_i.
  - Otherwise implement $Y_i = 0$ and charge nothing.
- Dominant strategy for individual i: Truthfully report V_i.
- This mechanism reveals the entire response function $x \to Y_i(x)$!

  Becker, DeGroot, and Marschak [1964].

Pricing using adaptive BDM mechanism

- BDM mechanism makes observability independent of chosen X_i.
- $\implies$ pure online learning problem (no exploration motive).
- Can use standard online-learning approaches:
  - Bayesian: Choose X_i to maximize $E[R_i(x)|W_i, V_1, \ldots, V_{i-1}]$.
  - Adversarial: Choose X_i with probability proportional to $\exp\left(\eta \cdot \sum_{i'<i} R_{i'}(x)\right)$.
- Generalization: Reserve price setting in auctions.

  Nedelec, Calauzènes, El Karoui, Perchet, et al. [2022].

Eliciting externalities (work in progress)

- Suppose i's utility depends on actions of *other* individuals.
- Such externalities - by definition - don't affect observable behavior.
- How can we possibly learn about them?
- VCG style mechanism:
  - Let individuals choose a Pigou tax affecting others.
  - At quadratic cost for themselves.

Binary choice with externalities

- Utility
$$U_i(Y_i, T_i) = \underbrace{Y_i \cdot (V_i - T)}_{\text{Utility of own action}} - \underbrace{B_i \cdot \bar{Y}}_{\text{Externality}} - \underbrace{\Delta_i(Y_i, T_i)}_{\text{Cost of taxing others}} .$$
- As before: Action Y_i, price T, w.t.p. V_i.
- Denote $\bar{Y} = \sum_i Y_i$.
- Additionally:
  - Externality $B_i \cdot \bar{Y}$ of action of others.
  - Choose price increment for other people, where
  $$T = \sum_i T_i.$$
  - Incur cost $\Delta_i(T_i)$ (to be specified).

Approximate VCG mechanism

- Denote $\bar{Y}(t) = \sum_i Y_i(t)$.
- Assume $\bar{Y}(t) \approx \bar{Y}(T) + \bar{Y}' \cdot (t - T)$.
- Set $\Delta_i(T_i, Y_i) = \frac{\bar{Y}'}{2} \cdot T_i^2 + Y_i \cdot T_i$, and let $T_{-i} = T - T_i$ so that

$$U_i(Y_i, T_i) = Y_i \cdot (V_i - T_{-i}) - B_i \cdot \bar{Y}(T) - \frac{\bar{Y}'}{2} \cdot T_i^2.$$

- Individual optimality condition for T_i:

$$B_i \cdot \bar{Y}' = \bar{Y}' \cdot T_i,$$

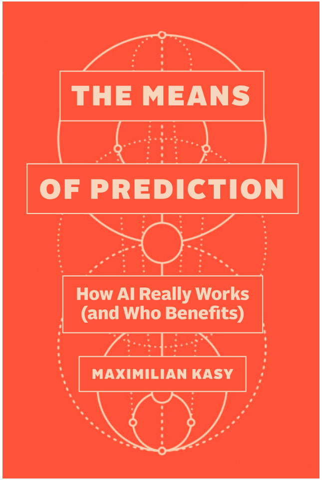
  so that $T_i = B_i$
- $\implies$ truthful revelation of externalities.

Adaptive Pigou taxation

- Repeated experiment in batches g.
- Externalities only within batches.
- Maintain running estimate of $\bar{Y}'$.
- For each batch g:
  1. Ask individuals to report $T_i = B_i$.
  2. Fix price $T = \sum_i T_i$.
  3. Ask individuals to report willingness to pay V_i.
  4. Implement $Y_i = \mathbf{1}(V_i \geq T_{-i})$ and payment

$$Y_i \cdot T_{-i} - \frac{\bar{Y}'}{2} \cdot T_i^2.$$

On a separate note



Key argument of the book

1. AI is automated decision-making using *optimization*.
2. Key issue: Who gets to pick the *objectives* that AI optimizes? (Not: Did the AI fail to optimize?)
3. Power flows from control of AI *inputs*: data, compute, expertise, energy.
4. We need *democratic control* of AI objectives by those affected by AI decisions.

Thank you!