Machine Learning Department
Carnegie Mellon University
School of Computer Science

Representation Learning for Sample-Efficient CATE Estimation

Maitreyi Swaroop, Bryan Wilder
Carnegie Mellon University

Motivation: *CATE is crucial for targeted interventions*

$T=0$ $T=1$

Problem: RCTs are small and expensive + long term outcomes are not observed till much later

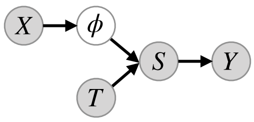
Silver lining: There is plenty of rich observational (administrative) data
The catch: It is noisy, and often lacks a clear treatment assignment

Additionally: Administrative data measures many outcomes beyond the primary target (Y).

- **Secondary outcomes** (S) act as noisy indicators of latent factors driving treatment heterogeneity.

Main idea: Learn a representation of covariates (X) capturing shared structure with S on large observational data
→ Use this lower-dimensional $\phi(X)$ to estimate CATE on fewer experimental samples
→ Substantially improved sample efficiency!

Method Overview



STAGE 1: Observational Data

Learn representation $\phi(X)$ **from** (X, S, Y) → **Fit a prediction head** $h : (\phi(X) \times S) \rightarrow Y$

STAGE 2: Experimental Data

$T=0$ $T=1$

Estimate CATE as

$$\tau(x) = \mathbb{E}\left[h(\phi(x), S) \mid \phi(x), T=1, P=E\right] - \mathbb{E}\left[h(\phi(x), S) \mid \phi(x), T=0, P=E\right]$$

Assumptions:

- **Ignorability/Unconfoundedness:** $(Y(1), Y(0)) \perp T \mid X, P=E$
- **Overlap:** $0 < \Pr[T=1 \mid X=x] < 1 \, \forall x$
- **Surrogacy:** $T \perp Y \mid S, X, P=E$
- **Comparability:** $P \perp Y \mid S, X$
- **Sufficiency:**
  i. $Y \perp X \mid \phi(X), S, P=O$
  ii. $S \perp X \mid \phi(X), T, P=E$

Objective function used to obtain ϕ

$$\phi^* = \arg\max_{\phi} \quad I(Y; \phi(X) \mid S) + \lambda \cdot I(S; \phi(X))$$

where λ is a hyper-parameter

Main result: Under assumptions 1-5, the CATE is identified using ϕ obtained via the above optimization.

Empirical Studies

Dataset

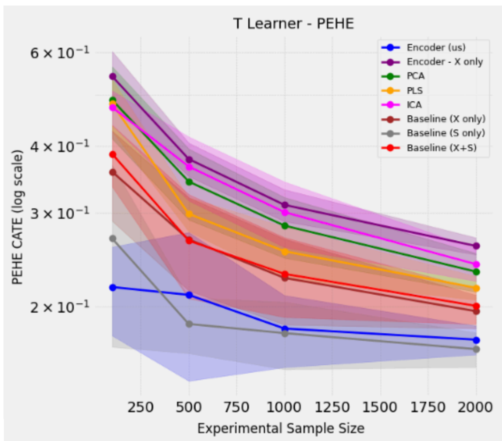
- RCT of free public transit fares in Allegheny County.
- Primary outcome (Y): total earnings post enrollment in RCT
- Secondary outcomes (S):
  1. Employment Status (binary indicator of any paid employment post RCT)
  2. Healthcare utilization (# of days with non-ER outpatient physical health visits)

* All outcomes are reported for the first 6 calendar quarters post RCT.

Setup

- Keeping the size of the observational sample fixed, we vary the size of the experimental sample
- Fit a CATE estimator (T-learner) on the experimental sample
- Baselines:
  1. Directly using the higher dimensional covariates
  2. Dimension reduction of covariates using PCA

PEHE of estimated CATE vs sample size



Observation: Our method (encoder) yields improved sample efficiency in CATE estimation.