Representation Learning for Treatment Effect Estimation from Observational Data

Liuyi Yao[1], Sheng Li[2], Yaliang Li[3], Mengdi Huai[1], Jing Gao[1], Aidong Zhang[1]

[1] University at Buffalo, [2] University of Georgia, [3] Tencent Medical AI Lab

Problem Setting

Treated group X_T
(with treatment t = 1)

Control group X_C
(with treatment t = 0)

- **Individual treatment effect (ITE)** of unit i:

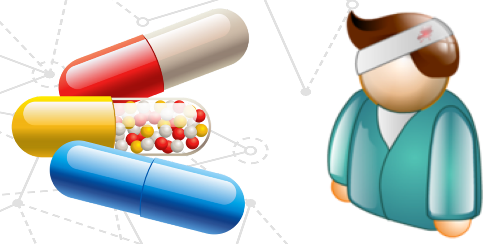
$$\underbrace{ITE_i}_{} = \underbrace{Y_{t=1}(x_i)}_{\text{Treated outcome}} - \underbrace{Y_{t=0}(x_i)}_{\text{Control outcome}}$$

- Estimating ITE can **help** decision making:

Policy Decision:

Who will benefit from job training?

Healthcare:



Which medicine?

- Why from the **observational data**?

Randomized Controlled Trials (RCTs) → Need to control the treatment assignment

Expensive ☹ | Time consuming ☹ | Ethical issues… ☹

Observational Data → Natural observations

Easy to access 😄 | Large amount of Data 😄

Challenges

- **Missing counterfactuals**

$$ITE_i = \boxed{Y_{t=1}(x_i)} - \boxed{Y_{t=0}(x_i)}$$

One is always missing in the data;
The missing one is the **counterfactual.**

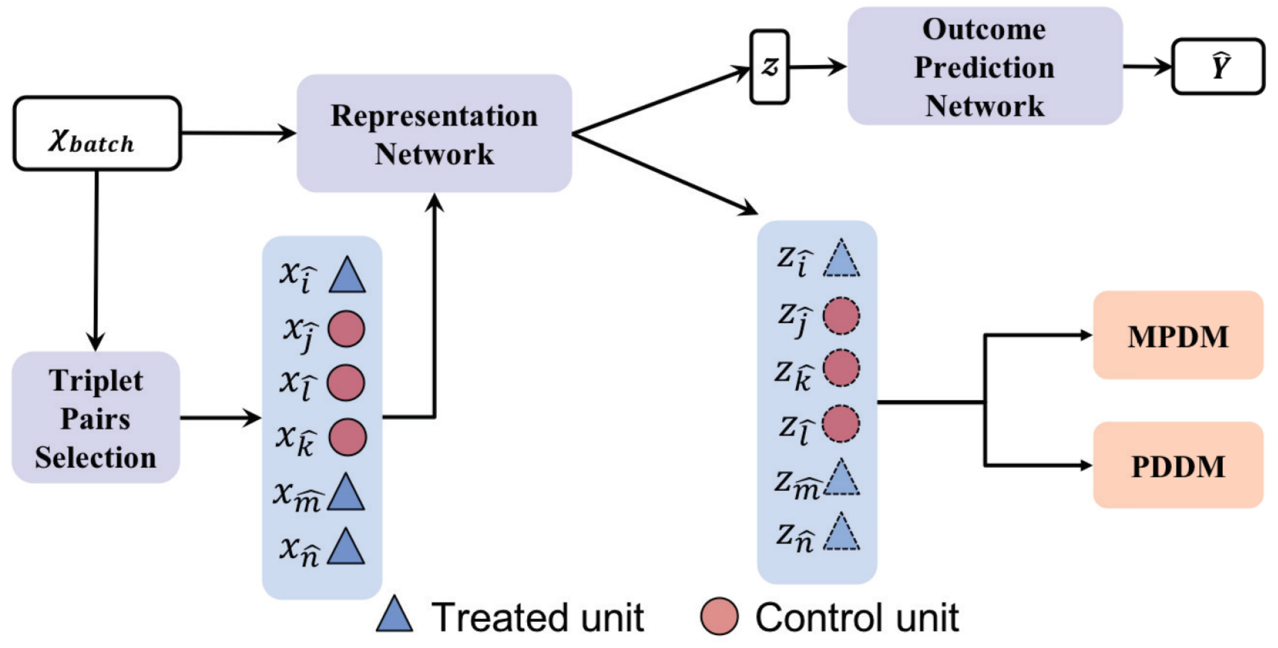
- **Selection bias**
  - Distributions of control & treated groups:

$$P(X_c) \neq P(X_T)$$

  - Counterfactual inference is more difficult.
  - Provided data: $(\mathbf{X_c}, \mathbf{Y_{t=0}})$, $(\mathbf{X_T}, \mathbf{Y_{t=1}})$
    - $(\mathbf{X_T}, \mathbf{Y_{t=1}}) \xrightarrow{\text{infer}} \mathbf{Y_{t=1}} | \mathbf{X_c}$
    - $(\mathbf{X_c}, \mathbf{Y_{t=0}}) \xrightarrow{\text{infer}} \mathbf{Y_{t=0}} | \mathbf{X_T}$

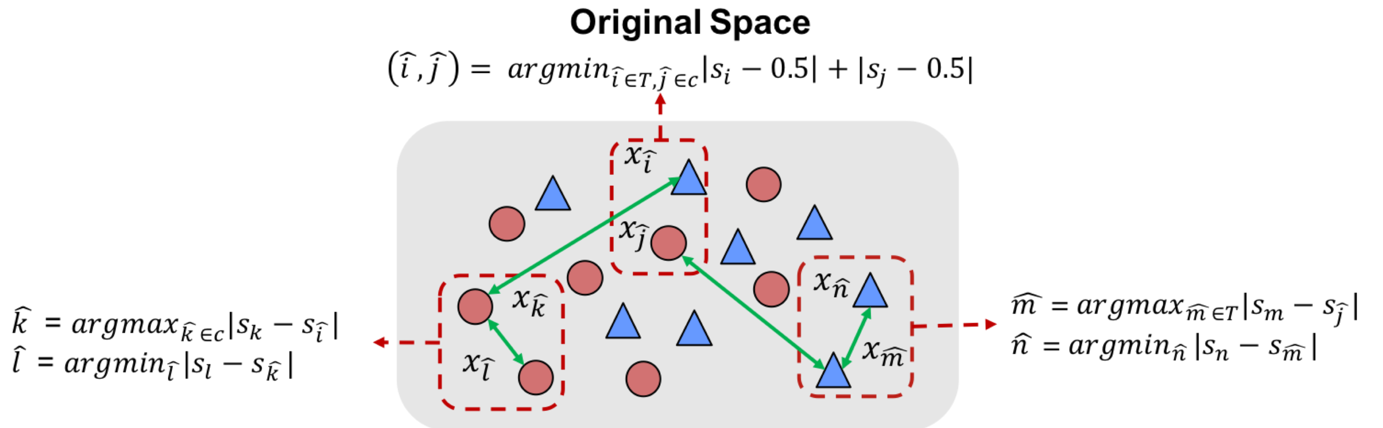
Our Framework

Local **s**imilarity preserved **i**ndividual **t**reatment **e**ffect estimation (**SITE**):



- The latent space Z should encode:
  - The balanced distributions;
  - The similarity information in X. ⟶ *Existing works ignore this.*
- SITE takes input units in a mini-batch fashion, which improves efficiency.

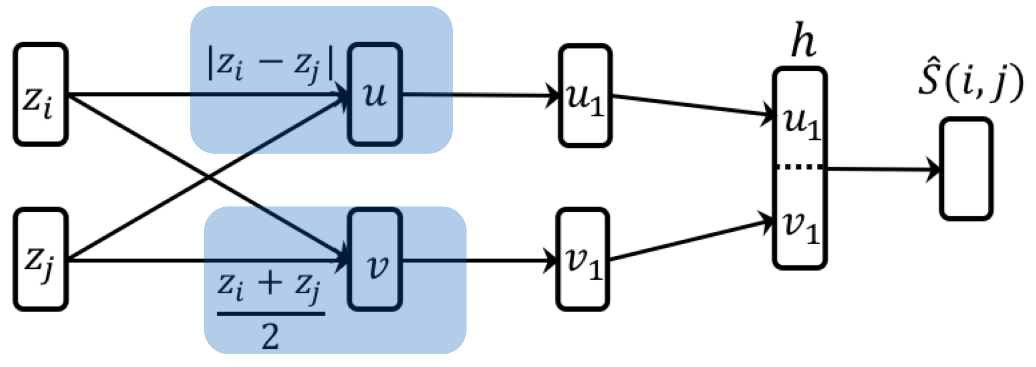
Triplet Pair Selection



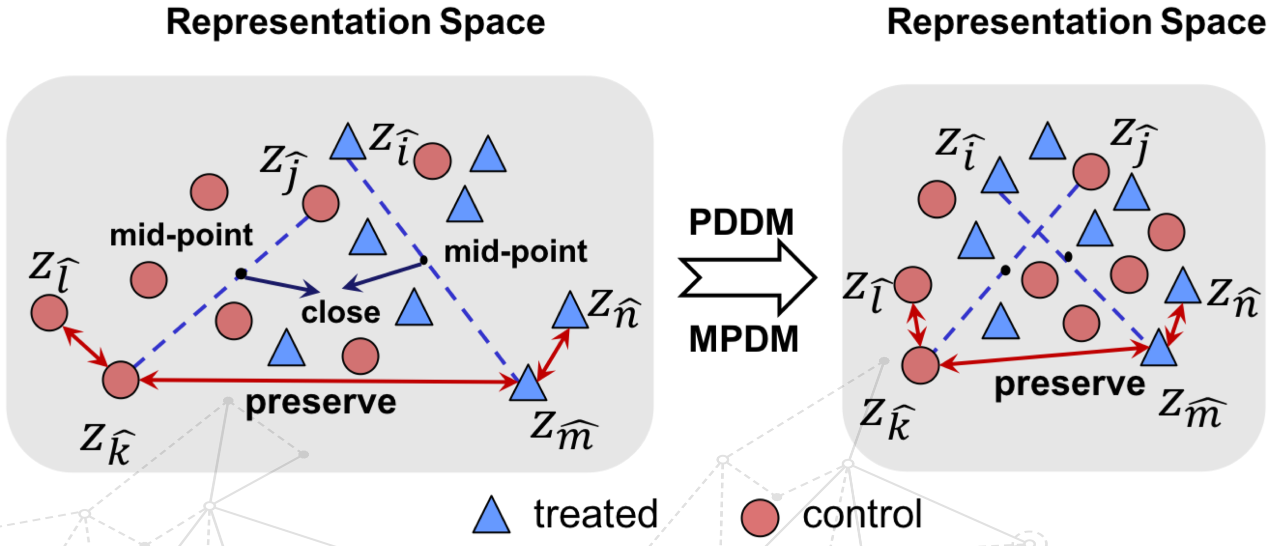
- s_*: propensity score (probability of a unit in treated group). ⟶ Can reflect the relative location of units in the original space.

PDDM & MPDM

- Position-Dependent Deep Metric (**PDDM**):
  - **Preserve the similarity.**



- Middle Point Distance Minimization (**MPDM**):
  - MPDM makes the middle point of $(z_{\hat{i}}, z_{\hat{m}})$ close to the middle point of $(z_{\hat{j}}, z_{\hat{k}})$.
  - **Balance the distribution in the latent space.**



Experiment Setting

IHDP Dataset:

- **Treatment:** specialist home visits;
- **Outcome:** Infants' cognitive test scores;
- **Pre-treatment covariates:** 25 covariates measuring aspects of children and their mothers;
- **Performance measurement:** Precision in Estimation of Heterogeneous Effect (ε_{PEHE}).

Jobs Dataset:

- **Treatment:** Job training;
- **Outcome:** employment status after training;
- **Pre-treatment covariates:** 8 covariates, such as age, education, ethnicity, previous earnings;
- **Performance measurement:** policy risk (R_{pol}).

Twins Dataset:

- **Treatment:** being the heavier one in twin;
- **Outcome:** one year mortality;
- **Pre-treatment covariates:** 40 covariates measuring aspects of pregnancy;
- **Performance measurement:** AUC.

Experimental Results

- Results on three datasets:

	IHDP ($\mathcal{E}_{\text{PEHE}}$)		**Jobs** ($\mathcal{R}_{\text{pol}}$)		**Twins** (AUC)	
Method	Within-sample	Out-of-sample	Within-sample	Out-of-sample	Within-sample	Out-of-sample
OLS/LR$_1$	10.761 ± 4.350	7.345 ± 2.914	0.310 ± 0.017	0.279 ± 0.067	0.660 ± 0.005	0.500 ± 0.028
OLS/LR$_2$	10.280 ± 3.794	5.245 ± 0.986	0.228 ± 0.012	0.733 ± 0.103	0.660 ± 0.004	0.500 ± 0.016
HSIC-NNM	14.384 ± 5.377	14.230 ± 5.566	0.282 ± 0.013	0.286 ± 0.080	0.762 ± 0.011	0.501 ± 0.017
PSM	7.490 ± 2.486	7.324 ± 2.230	0.284 ± 0.034	0.416 ± 0.313	0.500 ± 0.003	0.506 ± 0.011
k-NN	7.386 ± 2.623	7.443 ± 2.530	0.295 ± 0.026	0.416 ± 0.313	0.609 ± 0.010	0.492 ± 0.012
BNN	3.827 ± 2.044	4.874 ± 2.850	0.232 ± 0.008	0.240 ± 0.012	0.690 ± 0.008	0.676 ± 0.008
TARNet	0.729 ± 0.088	1.342 ± 0.597	0.228 ± 0.004	0.234 ± 0.012	0.849 ± 0.002	0.840 ± 0.006
CFR-MMD	0.663 ± 0.068	1.202 ± 0.550	**0.213 ± 0.006**	0.231 ± 0.009	0.852 ± 0.001	0.840 ± 0.006
CFR-WASS	0.649 ± 0.089	1.152 ± 0.527	0.225 ± 0.004	0.225 ± 0.010	0.850 ± 0.002	0.842 ± 0.005
SITE	**0.604 ± 0.093**	**0.656 ± 0.108**	0.224 ± 0.0004	**0.219 ± 0.009**	**0.862 ± 0.002**	**0.853 ± 0.006**

- **Code** of SITE: https://github.com/Osier-Yi/SITE