Counterfactual Prediction for Outcome-Oriented Treatments

Hao Zou
Tsinghua University

Bo Li
Tsinghua University

Jiangang Han
Meituan

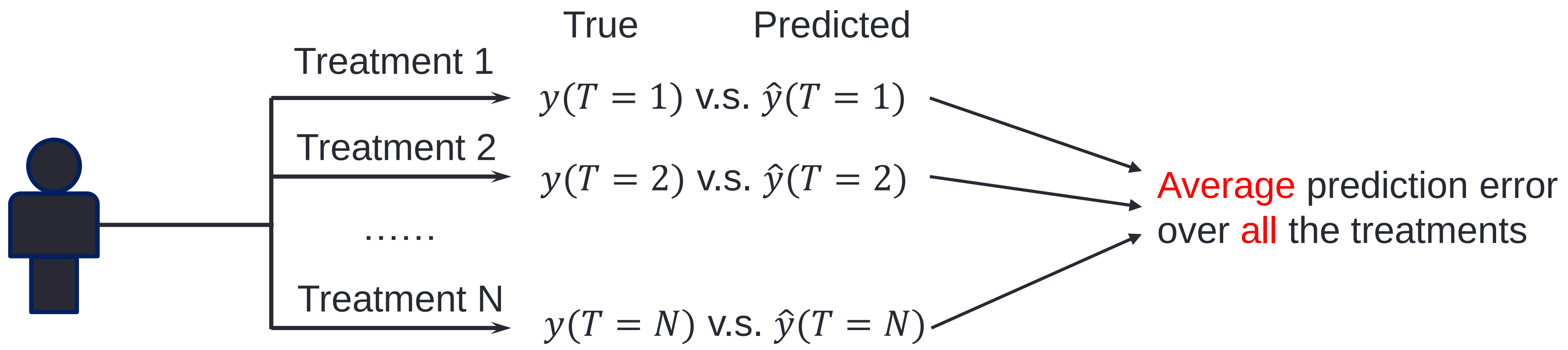
Shuiping Chen
Meituan

Xuetao Ding
Meituan

Peng Cui
Tsinghua University

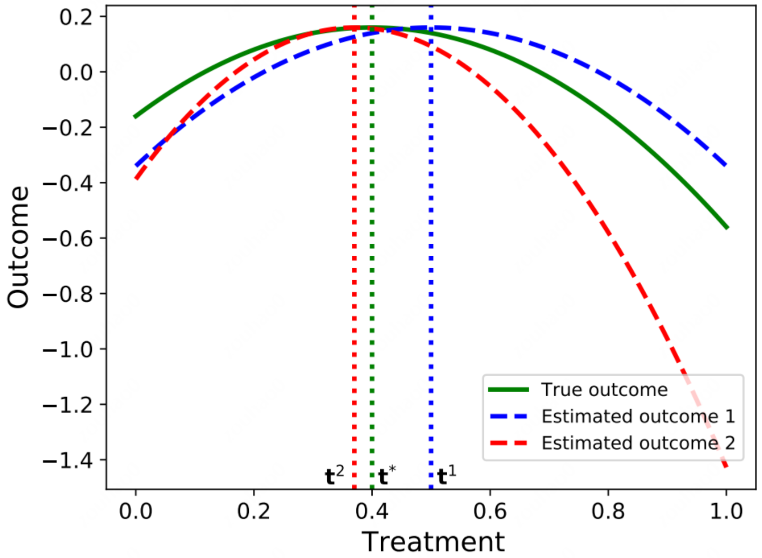
Background

- Large amounts of counterfactual prediction works exists
  - For binary treatments, categorical treatments, multi-dimensional treatments
  - Under static setting and time-series setting.
- The target of counterfactual prediction
  - PEHE (Precision in Estimating Heterogenous Treatment Effect) for binary treatment
  - Average outcome prediction error for more complex treatment



Background

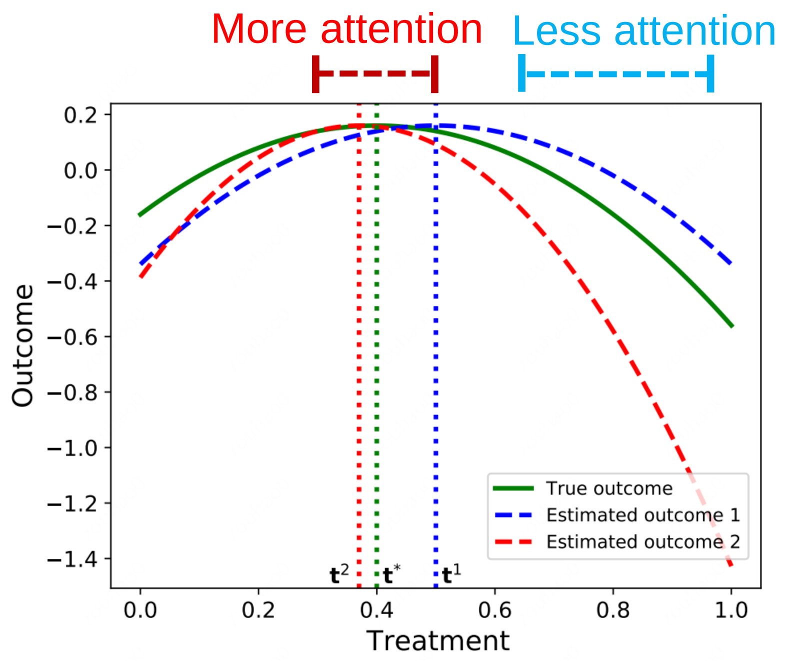
- More Accurate Prediction≠ Better Decision Making[1]
  - For example, green line represents the true outcome curve, red and blue lines represent two estimated outcome curves.
  - The blue estimated outcome curve ⟶ smaller prediction error
  - Blue optimal treatment t^1 is worse than red optimal treatment t^2



[1] Fernandez-Lorıa, C. and Provost, F. Causal decision making and causal effect estimation are not the same. . . and why it matters. INFORMS Journal on Data Science, 2022.

Background

- For decision-making, not all treatments are equally important
  - When selecting movies, people pay more attention to popular movies.
  - When hiring an employee, interviewer concentrate on the competitive candidates
  - …
- We focus more on Outcome-Oriented Treatments for counterfactual prediction



Problem Formulation

- We consider the continuous treatment setting.
- **Target**: Learning counterfactual prediction model from observational dataset
- **Observational dataset**: $\{(x_i, t_i, y_i)\}_{i=1,2,3,..,n}$, where n is the sample size
  - $x_i \in \mathcal{X}$ is the confounder variables.
  - $t_i \in \mathcal{T} = [a, b]$ is the continuous treatment
  - $y_i \in \mathbb{R}$ is the corresponding outcome
- **Evaluation**: Treatment selection regret for model $f: \mathcal{X} \times \mathcal{T} \longrightarrow \mathbb{R}$

$$Regret(f) = \mathbb{E}_{\mathbf{X}}\left[Y_{\mathbf{X}}(\rho^*(\mathbf{X})) - Y_{\mathbf{X}}(\rho^f(\mathbf{X}))\right]$$

$Y_X(\mathbf{t})$ is potential outcome function

$$\rho^*(\mathbf{X}) = \arg\max_{\mathbf{t}} Y_{\mathbf{X}}(\mathbf{t}),$$

$$\rho^f(\mathbf{X}) = \arg\max_{\mathbf{t}} f(\mathbf{X}, \mathbf{t})$$

Theoretical Analysis on Regret

- We can have the following upper bound of treatment selection regret

Proposition 4.1. *With the confounders* $\mathbf{X}$, *treatments* $\mathbf{t}$, *potential outcome function* $Y_{\mathbf{X}}(t)$ *defined as above, the treatment selection regret (i.e. Equation 1) of counterfactual prediction model* f *satisfies the following inequality:*

$$Regret(f) \leq \sqrt{\mathbb{E}_{\mathbf{X}}[(Y_{\mathbf{X}}(\rho^f(\mathbf{X})) - f(\mathbf{X}, \rho^f(\mathbf{X})))^2]} + \sqrt{\mathbb{E}_{\mathbf{X}}[(Y_{\mathbf{X}}(\rho^*(\mathbf{X})) - f(\mathbf{X}, \rho^*(\mathbf{X})))^2]} \quad (4)$$

Approximate (first term):

$$\frac{1}{n}\sum_{i=1}^{n} \frac{K\left((\rho^f(\mathbf{x}_i) - t_i)/\tau\right)}{\tau p(t_i|\mathbf{x}_i)}(y_i - f(\mathbf{x}_i, t_i))^2 \triangleq \mathcal{A}(f)$$

Upper bound (second term):

$$\frac{1}{n}\sum_{i=1}^{n} \frac{(y_i - f(\mathbf{x}_i, t_i))^2}{(b-a)p(t_i|\mathbf{x}_i)} + L \cdot \frac{b-a}{2} \triangleq \mathcal{B}(f)$$

Objective Function

- We can obtain the upper bound $\sqrt{\mathcal{A}(f)} + \sqrt{\mathcal{B}(f)}$ of regret.
- For the stability of training process, we optimize $\gamma\mathcal{A}(f) + \mathcal{B}(f)$ instead.
- Therefore, the final loss function for model with parameter θ is:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \frac{1+\lambda K\big((\rho^{f_\theta}(\mathbf{x}_i)-t_i)/\tau\big)}{(b-a)p(t_i|\mathbf{x}_i)} \cdot (y_i - f_\theta(\mathbf{x}_i, t_i))^2$$

λ and τ is hyper-parameters.

Proposition 4.5. *Assuming the function is parameterized by θ, that is f_θ, and the functions $\mathcal{A}(f_\theta)$ and $\mathcal{B}(f_\theta)$ are differentiable and strictly convex on θ, θ^* is the global minimum point of $\sqrt{\mathcal{A}(f_\theta)} + \sqrt{\mathcal{B}(f_\theta)}$, then there exists $\gamma \in \mathbb{R}^+$ such that*

$$\theta^* = \arg\min_{\theta} \gamma\mathcal{A}(f_\theta) + \mathcal{B}(f_\theta) \quad (9)$$

Implementation

- The components of our algorithm are implemented as following:
- Inverse propensity score:
  - We label $\{(x_i, t_i)\}_{1 \leq i \leq n}$ with positive label (L=1) and label $\{(x_i, t_i')\}_{1 \leq i \leq n}, t_i' \sim Unif(a, b)$ with negative label (L=0). After training a classifier $\hat{p}(L|x, t)$ on these samples, we have
  - $$\frac{1}{\hat{p}(\mathbf{t}|\mathbf{X})} = \frac{(b-a)\hat{p}(L=0|\mathbf{X},\mathbf{t})}{\hat{p}(L=1|\mathbf{X},\mathbf{t})}$$
- Outcome-oriented sample re-weighting:
  - In the first stage, we train the model with sample weights $w_i^{(0)} = \frac{1}{(b-a)\hat{p}(t_i|\mathbf{x}_i)}$
  - In the second stage, we train the model for m rounds. For j^{th} round, we train the model with sample weights $w_i^{(j)}$ and obtain the model $f^{(j)}$

$$w_i^{(j)} = \frac{1 + \lambda K\left((\rho^{f_\theta^{(j-1)}}(\mathbf{x}_j) - t_j)/\tau\right)}{(b-a)\hat{p}(t_j|\mathbf{x}_j)}$$

- Outcome prediction:
  - The loss function for training model at j^{th} round is $\mathcal{L}^{(j)} = \frac{1}{n}\sum_{i=1}^{n} w_i^{(j)} \cdot (f_\theta^{(j)}(\mathbf{x}_i, t_i) - y_i)^2$

Empirical Results

- We compare our method with some baselines,
  - including SCIGAN, RMNet, IPS-BanditNet…
- Evaluation metric: Treatment selection regret
  - Within-sample setting: $Regret_{in} = \frac{1}{n}\sum_{i=1}^{n}\left(Y_{\mathbf{x}_i}(\rho^*(\mathbf{x}_i)) - Y_{\mathbf{x}_i}(\rho^f(\mathbf{x}_i))\right)$
  - Out-of-sample setting: Average over new samples $\{\mathbf{x}_i^{tes}\}_{1\leq i\leq n_{tes}}$
- Experiments on both synthetic-datasets and semi-synthetic datasets
- The pseudo-optimal treatment of model f

$$\rho^f(\mathbf{X}) = \arg\max_{\mathbf{t}\in\{a, a+\frac{b-a}{q-1}, \ldots, b\}} f(\mathbf{X}, \mathbf{t})$$

We set $q = 1001$

Generating Synthetic Datasets

- We generate the datasets as following
  - Generate confounders $\mathbf{X} = (x_1, x_2, ..., x_d)$ $x_j = |u_j|$ $u_j \overset{iid}{\sim} \mathcal{N}(0,1)$
  - Generate two constant vectors $\mathbf{v}_1 \in \mathbb{R}^{d\times 1}$ and $\mathbf{v}_2 \in \mathbb{R}^{d\times 1}$
  - We define the outcome generation mechanism as $Y_{\mathbf{X}}(\mathbf{t}) = g(\mathbf{X},\mathbf{t}) \cdot \mathbf{t}$ similar to gross merchandise volume (GMV) in marketing.
  - Mimicking the demand curve in marketing[1], we set three forms of $g(\mathbf{X},\mathbf{t})$
    - Linear: $g(\mathbf{X},\mathbf{t}) = max(-\mathbf{v}_2^{\top}\mathbf{X} \cdot \mathbf{t} + 1.8\mathbf{v}_1^{\top}\mathbf{X}, 0)$
    - Exponential: $g(\mathbf{X},\mathbf{t}) = e^{-\mathbf{v}_2^{\top}\mathbf{X}\cdot\mathbf{t}+\mathbf{v}_1^{\top}\mathbf{X}}$
    - Logit: $g(\mathbf{X},\mathbf{t}) = 2/(1 + e^{\mathbf{v}_2^{\top}\mathbf{X}\cdot\mathbf{t}-\mathbf{v}_1^{\top}\mathbf{X}})$
  - The treatments are sampled from Beta distribution $\frac{t_i}{r} \sim \mathrm{Beta}(\alpha, \beta)$

$$\beta = \frac{\alpha - 1}{\rho^*(\mathbf{x}_i)/2r} + 2 - \alpha$$

[1] Besbes, O. and Zeevi, A. On the (surprising) sufficiency of linear models for dynamic pricing with demand learning. Management Science, 61(4):723–739, 2015.

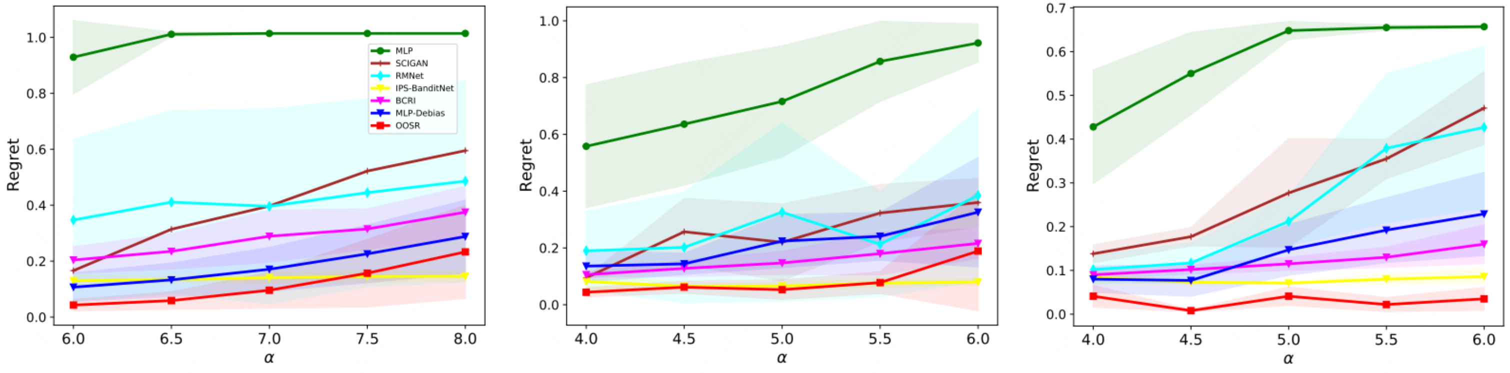
Results on Synthetic Datasets

- Varying sample size: (Part of results)

Linear setting: Fix the degree of selection bias $\alpha = 6.0$, varying the sample size n								
n	$n = 4000$		$n = 6000$		$n = 8000$		$n = 10000$	
Methods	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.
MLP	0.914±0.133	0.929±0.131	0.887±0.160	0.895±0.160	0.804±0.236	0.811±0.239	0.833±0.207	0.849±0.208
SCIGAN	0.156±0.002	0.166±0.002	0.140±0.002	0.146±0.003	0.126±0.002	0.132±0.002	0.130±0.003	0.136±0.002
RMNet	0.343±0.285	0.347±0.290	0.286±0.241	0.287±0.244	0.181±0.098	0.178±0.096	0.192±0.136	0.193±0.137
IPS-BanditNet	0.125±0.021	0.130±0.022	0.105±0.018	0.109±0.019	0.104±0.014	0.108±0.015	0.103±0.019	0.107±0.020
BCRI	0.199±0.046	0.204±0.047	0.172±0.035	0.175±0.035	0.150±0.026	0.154±0.027	0.137±0.015	0.139±0.014
MLP-Debias	0.100±0.048	0.107±0.051	0.081±0.057	0.083±0.058	0.074±0.047	0.073±0.047	0.053±0.029	0.055±0.030
OOSR	**0.040**±0.018	**0.043**±0.020	**0.034**±0.023	**0.046**±0.024	**0.020**±0.011	**0.037**±0.011	**0.015**±0.010	**0.016**±0.010

- Varying α

OOSR outperforms the baselines across different settings



(a) Linear setting (b) Exponential setting (c) Logit setting

Generating Semi-synthetic Datasets

- The confounder feature X is obtained from a real-world dataset TCGA[1].
- The outcome generation is:
  - Setting 1: $Y_{\mathbf{X}}(\mathbf{t}) = \mathbf{v}_1^\top \mathbf{X} + (12\mathbf{v}_2^\top \mathbf{X} - 2) \cdot \mathbf{t} - (12\mathbf{v}_3^\top \mathbf{X} - 2) \cdot \mathbf{t}^2$
  - Setting 2: $Y_{\mathbf{X}}(\mathbf{t}) = \mathbf{v}_1^\top \mathbf{X} + 12\mathbf{t} \cdot \left(\mathbf{t} - 0.75 \frac{\mathbf{v}_2^\top \mathbf{X}}{\mathbf{v}_3^\top \mathbf{X}}\right)^2$
- The treatments are sampled from Beta distribution $t_i \sim \mathrm{Beta}(\alpha, \beta)$ $\beta = \frac{\alpha - 1}{\rho^*(\mathbf{x}_i)/2} + 2 - \alpha$

Setting 1: Varying the degree of selection bias α

α	$\alpha = 6.0$		$\alpha = 6.5$		$\alpha = 7.0$		$\alpha = 7.5$	
Methods	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.
MLP	1.547±0.001	1.532±0.001	1.547±0.001	1.532±0.001	1.547±0.001	1.532±0.001	1.547±0.001	1.532±0.001
SCIGAN	0.251±0.006	0.254±0.006	0.387±0.008	0.392±0.008	0.551±0.010	0.556±0.009	0.785±0.013	0.792±0.013
RMNet	0.546±0.360	0.550±0.363	0.545±0.440	0.548±0.445	0.686±0.542	0.685±0.537	0.551±0.250	0.549±0.249
IPS-BanditNet	0.260±0.030	0.259±0.030	0.265±0.052	0.266±0.053	0.272±0.030	0.275±0.030	0.288±0.037	0.291±0.037
BCRI	0.091±0.063	0.093±0.061	0.121±0.088	0.124±0.090	0.186±0.039	0.187±0.038	0.502±0.176	0.499±0.171
MLP-Debias	0.040±0.014	0.039±0.014	0.202±0.071	0.204±0.071	0.276±0.083	0.278±0.086	0.346±0.090	0.352±0.093
OOSR	**0.016**±0.005	**0.015**±0.005	**0.096**±0.051	**0.097**±0.051	**0.125**±0.042	**0.127**±0.041	**0.187**±0.052	**0.190**±0.053

Setting 2: Varying the degree of selection bias α

α	$\alpha = 4.0$		$\alpha = 4.5$		$\alpha = 5.0$		$\alpha = 5.5$	
Methods	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.	Within-S.	Out-of-S.
MLP	0.100±0.064	0.098±0.058	0.210±0.058	0.195±0.054	0.192±0.068	0.182±0.063	0.279±0.073	0.266±0.071
SCIGAN	0.064±0.037	0.066±0.040	0.139±0.082	0.143±0.082	0.148±0.057	0.154±0.056	0.209±0.095	0.212±0.089
RMNet	0.154±0.064	0.159±0.065	0.145±0.068	0.149±0.070	0.165±0.129	0.169±0.128	0.189±0.080	0.192±0.075
IPS-BanditNet	0.509±0.044	0.496±0.045	0.491±0.033	0.473±0.034	0.580±0.125	0.569±0.133	0.623±0.156	0.608±0.155
BCRI	0.132±0.034	0.152±0.036	0.243±0.139	0.254±0.135	0.267±0.141	0.279±0.132	0.313±0.107	0.320±0.106
MLP-Debias	0.028±0.019	0.028±0.020	0.122±0.073	0.113±0.065	0.112±0.089	0.107±0.086	0.171±0.072	0.160±0.067
OOSR	**0.015**±0.014	**0.016**±0.015	**0.105**±0.079	**0.100**±0.071	**0.098**±0.096	**0.095**±0.093	**0.154**±0.066	**0.143**±0.062

[1] Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R. M., Ozenberger, B. A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J. M. The cancer genome atlas pan-cancer analysis project. Nature Genetics, 45:1113– 1120, 2013.

Conclusion

We theoretically analyze that decision-making performance is related to outcome prediction on true/pseudo-optimal treatments

We propose Outcome-Oriented Sample Re-weighting (OOSR) method to strengthen the prediction on outcome-oriented treatment region.

Experimental results on synthetic datasets and semi-synthetic datasets show the effectiveness of OOSR.

Thank you!