Causal Inference in Complex Environments

Kun Kuang

Zhejiang University

https://kunkuang.github.io/

An example of decision making

- Does predictive models guide decision making?
- System changes algorithm from A to B at some point.
- Is the new algorithm B better?
- Say algorithm that provides promotion or discount link to different customers

Algorithm A

Algorithm B

An example of decision making

- Measure success rate (SR)

Old Algorithm (A)	New Algorithm (B)
50/1000 **(5%)**	54/1000 **(5.4%)**

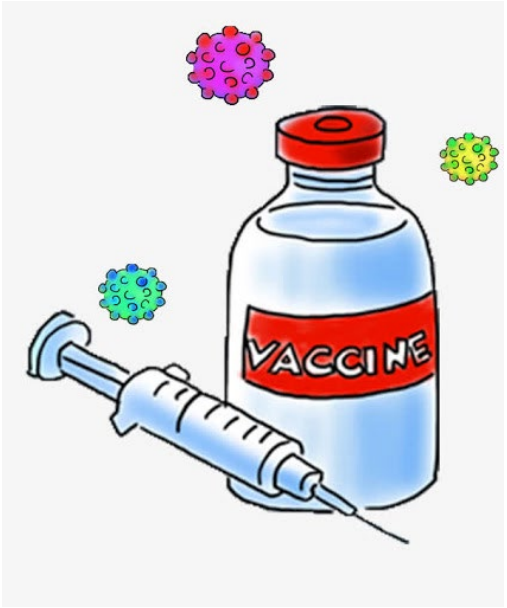
New algorithm increases overall success rate, so it is better?

	Old Algorithm (A)	New Algorithm (B)
Low-income Users	10/400 **(2.5%)**	4/200 **(2%)**
High-income Users	40/600 **(6.6%)**	50/800 **(6.2%)**
Overall	50/1000 **(5%)**	54/1000 **(5.4%)**

Which is better?

Decision Making with Causality

- Causal Effect Estimation is necessary for decision making!



Causal effect estimation plays an important role on decision making!

A practical definition

Definition: T causes Y if and only if

changing T leads to a change in Y,

keep everything else constant.

Causal effect is defined as the magnitude by which Y is changed by a unit change in T.

Two key points: changing T, keeping everything else constant

**Interventionist* definition [http://plato.stanford.edu/entries/causation-mani/]

Problem of Treatment Effect Estimation

- Treatment Variable: $T = 1$ or $T = 0$
- Potential Outcome: $Y(T = 1)$ and $Y(T = 0)$
- Individual Treatment Effect (ITE)

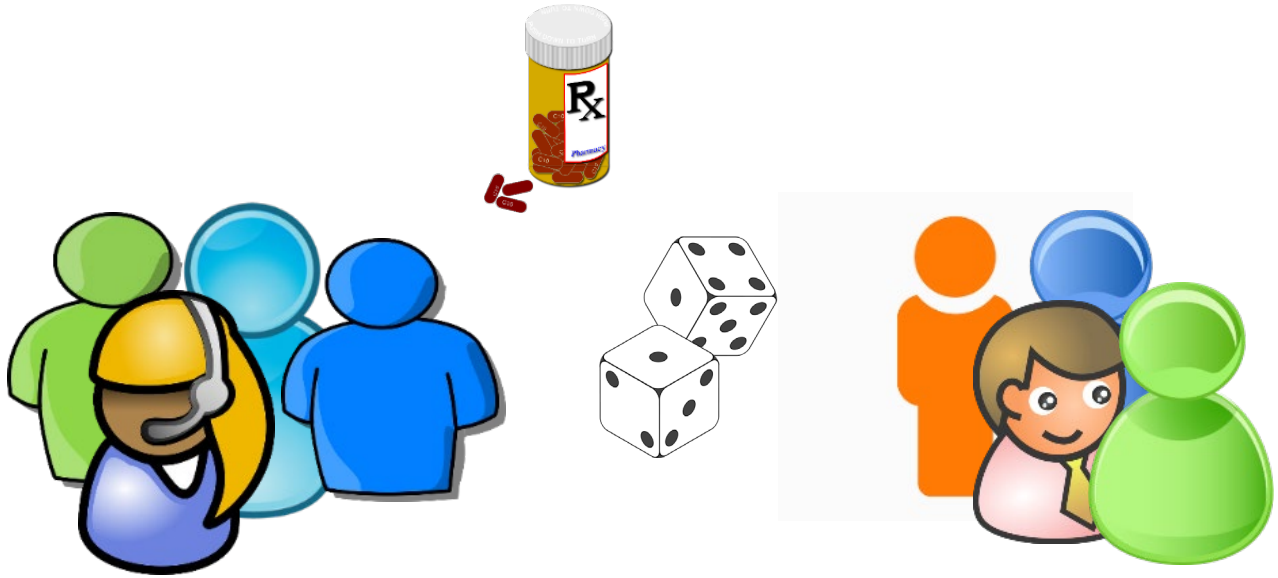
$$ITE(i) = Y_i(T_i = 1) - Y_i(T_i = 0)$$

- Average Treatment Effect (ATE):

$$ATE = E[Y(T = 1) - Y(T = 0)]$$

Counterfactual problem: $Y(T = 1)$ or $Y(T = 0)$

Randomized Experiments are the "Gold Standard"



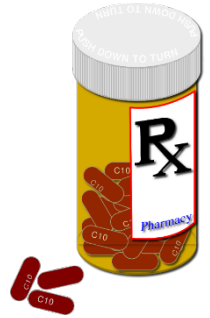
- Drawbacks of randomized experiments:
  - Cost
  - Unethical

- Two key points: changing T, keeping everything else constant

Causal Inference with Observational Data

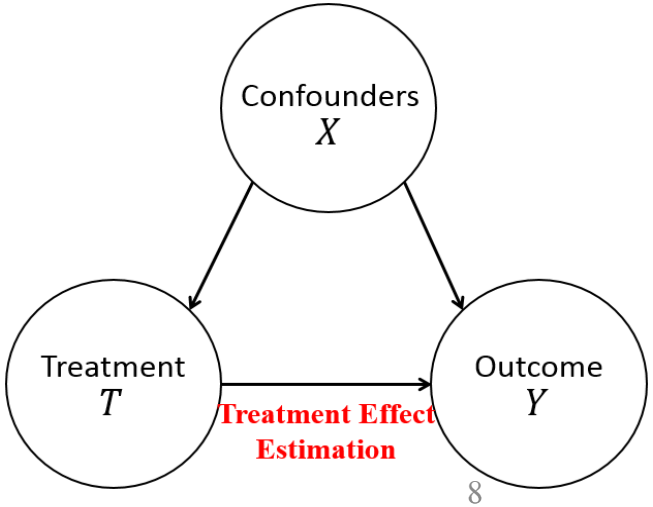
- Definition of ATE: $ATE = E[Y(T=1) - Y(T=0)]$
- In observational data, we have units with different T:

$$E[Y(T=1)] \quad \text{and} \quad E[Y(T=0)]$$



- Can we estimate ATE by directly comparing the average outcome between groups with T=1 and T=0?
  - No, because confounders X might not be constant

- Two key points:
  - Changing T (T=1 and T=0)
  - Keeping everything else (Confounder X) constant

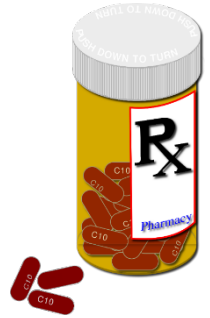


Causal Inference with Observational Data

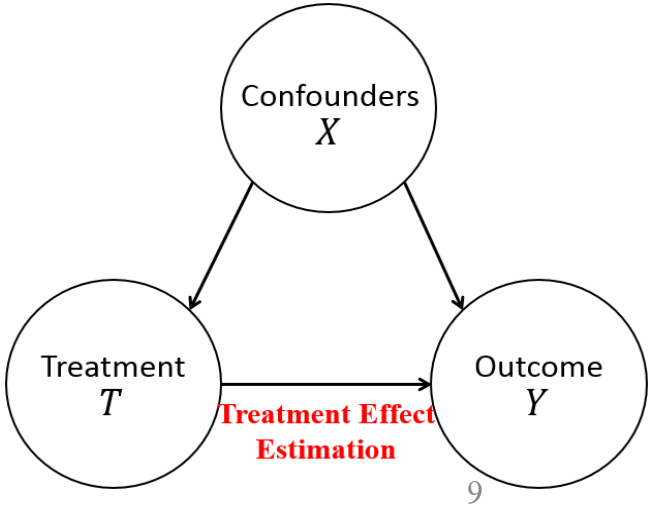
- Definition of ATE: $ATE = E[Y(T=1) - Y(T=0)]$
- In observational data, we have units with different T:

$$E[Y(T=1)] \quad \text{and} \quad E[Y(T=0)]$$

- Can we estimate ATE by directly comparing the average outcome between groups with T=1 and T=0?
  - No, because confounders X might not be constant



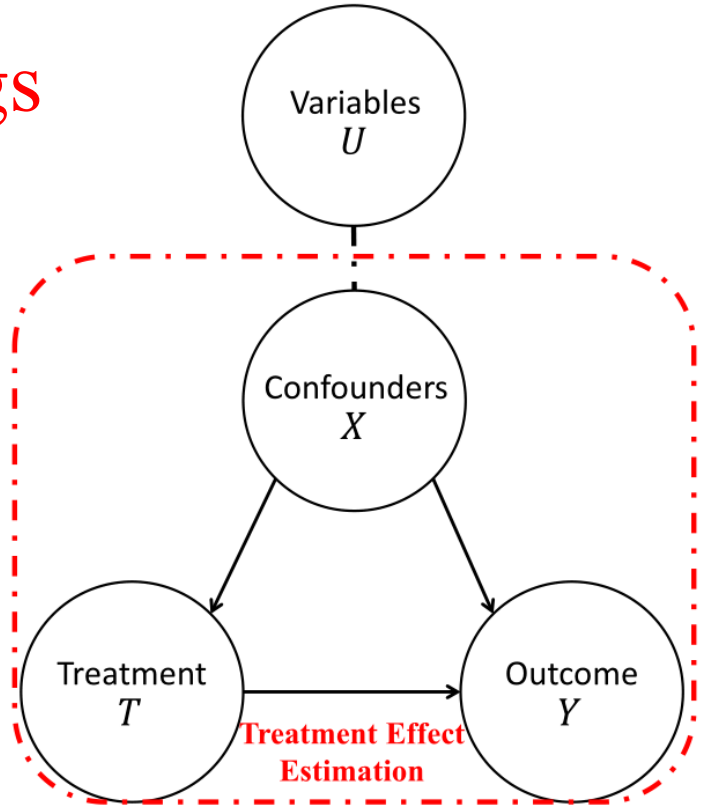
- Two key points:
  - Changing T (T=1 and T=0)



Balancing Confounders' Distribution

Related Work

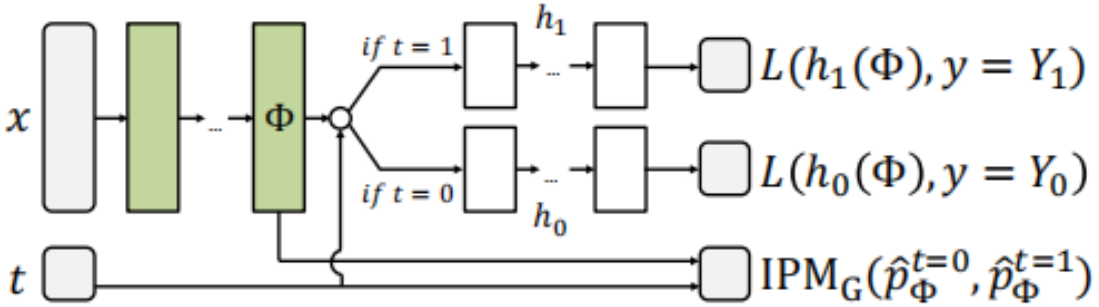
- Matching Methods
  - *Exactly Matching*, *Coarse Matching*
  - Poor performance in high dimensional settings
- Propensity Score based Methods
  - Propensity score $e(\mathbf{X}) = p(T = 1|\mathbf{X})$
  - *Matching*, *Weighting*, *Doubly Robust*
  - Treat all observed variables as confounders, and ignore the non-confounders
  - Mainly designed for binary treatment



(a) Previous Causal Framework.

Related Work

- Representation Learning based Methods
  - Similar representation between treatment groups.
  - Accurate prediction on factual/counterfactual outcome



$$\min_{\substack{h,\Phi \\ \|\Phi\|=1}} \frac{1}{n}\sum_{i=1}^{n} w_i \cdot L\left(h(\Phi(x_i), t_i), y_i\right) + \lambda \cdot \Re(h)$$
$$+ \alpha \cdot \mathrm{IPM}_G\left(\{\Phi(x_i)\}_{i:t_i=0}, \{\Phi(x_i)\}_{i:t_i=1}\right),$$
$$\text{with } w_i = \frac{t_i}{2u} + \frac{1-t_i}{2(1-u)}, \quad \text{where } u = \frac{1}{n}\sum_{i=1}^{n} t_i$$
and $\Re$ is a model complexity term.

- Confounder differentiation, binary treatment, might ignore confounders

Shalit U, Johansson F D, Sontag D. Estimating individual treatment effect: generalization bounds and algorithms. ICML 2017.

Standard Assumptions for Causal Inference

- **A1: Stable Unit Treatment Value (SUTVA):** The effect of treatment on a unit is independent of the treatment assignment of other units

$$P\left(Y_i \middle| T_i, T_j, X_i\right) = P(Y_i | T_i, X_i)$$

- **A2: Unconfounderness:** The distribution of treatment is independent of potential outcome when given the observed variables

$$T \perp \left(Y(0), Y(1)\right) | \, X$$

- **A3: Overlap:** Each unit has nonzero probability to receive either treatment status when given the observed variables

$$0 < P(T = 1 | X = x) < 1$$

New Challenges in Complex Environments

- Challenge 1: High dimensional variables, but NOT all variables are confounders.

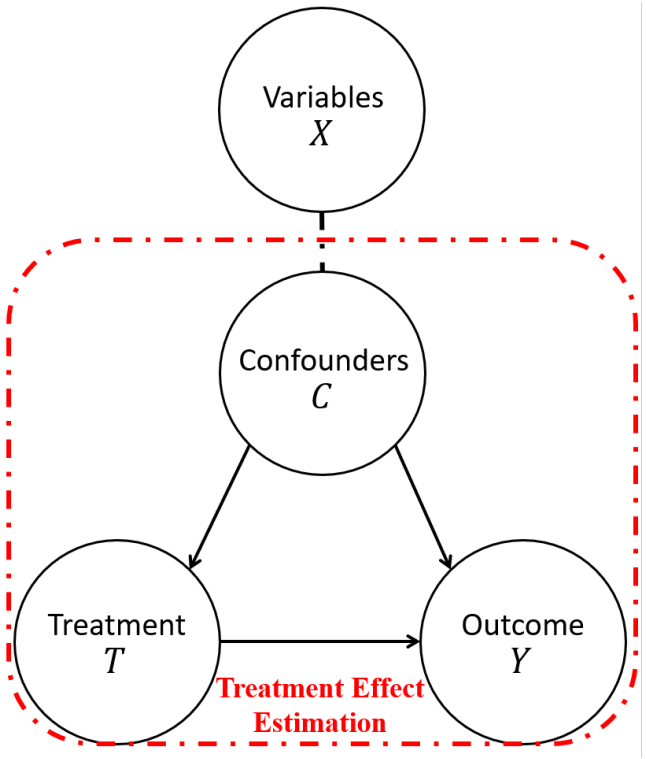
- Challenge 2: Unobserved confounders, NOT all confounders are observed. IV based method is a great approach for the problem, but limited to linear and requires pre-defined IV.

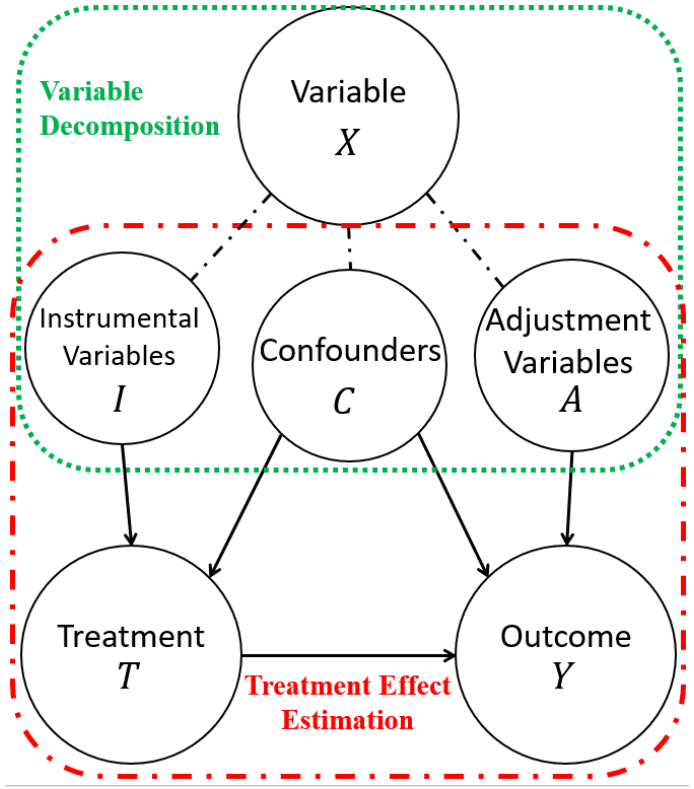
- Challenge 3: Complex Treatments without SUTVA assumption

Challenge 1: High dimensional variables

- With SUTVA, Unconfounderness, and Overlap Assumptions
- In complex environment, we may collect high-dimensional variables, including confounders and noisy variables
- But NOT all observed/collected variables are confounders, and including non-confounders might bring new bias
- How to automatically select the confounders for causal inference?

Learning Decomposed Representation for Counterfactual Inference





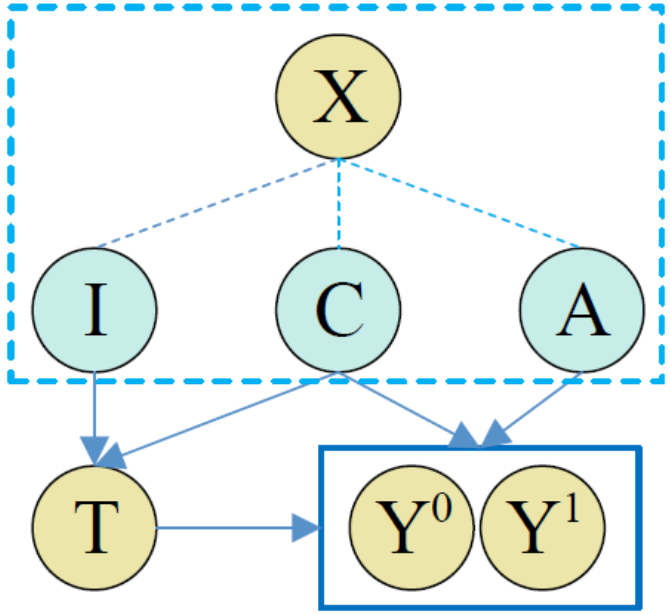
- All variables can be separated into 3 parts: IV, Confounders, Adjustment variables.
- Including IV will bring bias for causal inference.
- Including adjustment variables can help to reduce the variance.

Kuang K, Cui P, Li B, et al. Treatment effect estimation with data-driven variable decomposition [C]. AAAI, 2017
Kuang K, Cui P, et al. Data-Driven Variable Decomposition for Treatment Effect Estimation, TKDE, 2020
Wu A, Yuan J, Kuang K, et al. Learning decomposed representations for treatment effect estimation[J]. TKDE, 2022.

Learning Decomposed Representation for Counterfactual Inference

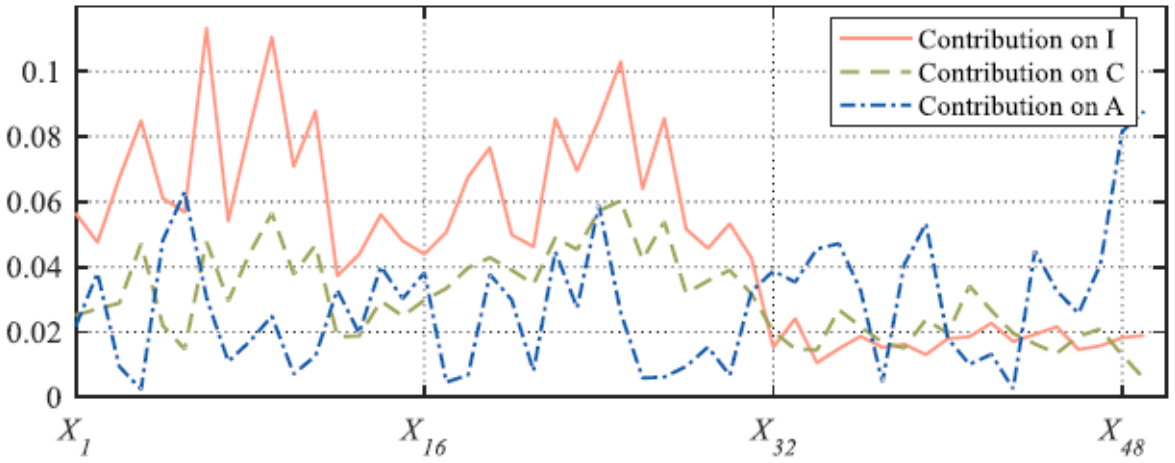
- Three decomposed representation networks
  - $I(X),\ C(X),\ A(X)$
- Three decomposition and balancing regularizers
  - Confounder identification: $A(X) \perp T, I(X) \perp Y \mid T$
  - Confounder balancing: $w \cdot C(X) \perp T$
- Two regression networks
  - $Y(T=1),\ Y(T=0)$
- Orthogonal Regularizer for Decomposition

$$\mathcal{L}_O = \bar{I}_W^T \cdot \bar{C}_W + \bar{C}_W^T \cdot \bar{A}_W + \bar{A}_W^T \cdot \bar{I}_W$$

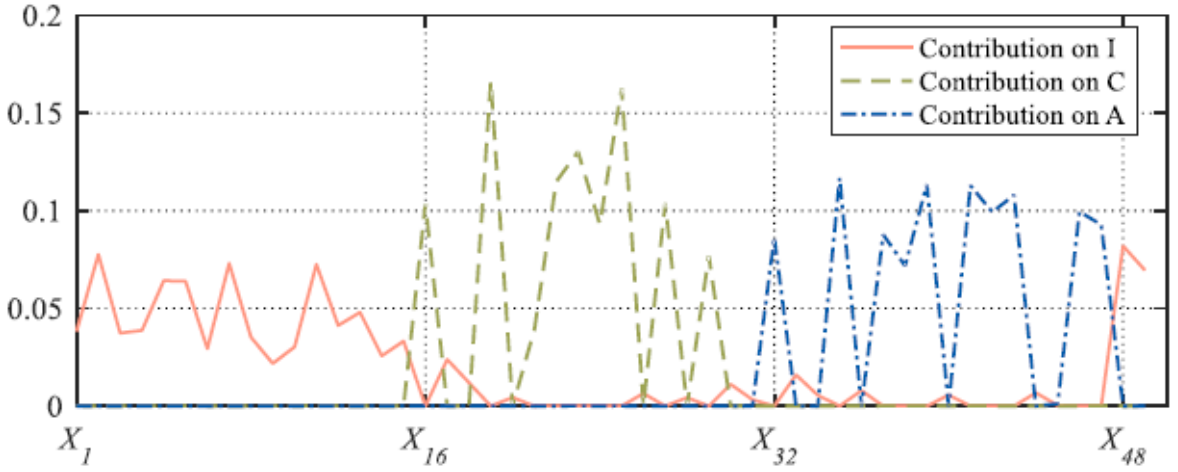


Wu A, Yuan J, Kuang K, et al. Learning decomposed representations for treatment effect estimation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

Learning Decomposed Representation for Counterfactual Inference



(a) DR-CFR in Syn_16_16_16_3000



(b) DeR-CFR in Syn_16_16_16_3000

Wu A, Yuan J, Kuang K, et al. Learning decomposed representations for treatment effect estimation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

Learning Decomposed Representation for Counterfactual Inference

Table 1: The results on IHDP.

IHDP				
Mean +/- Std	**Within-sample**		**Out-of-sample**	
Methods	PEHE	ϵ_{ATE}	PEHE	ϵ_{ATE}
CFR-MMD	0.702 +/- 0.037	0.284 +/- 0.036	0.795 +/- 0.078	0.309 +/- 0.039
CFR-WASS	0.702 +/- 0.034	0.306 +/- 0.040	0.798 +/- 0.088	0.325 +/- 0.045
CFR-ISW	0.598 +/- 0.028	0.210 +/- 0.028	0.715 +/- 0.102	0.218 +/- 0.031
SITE	0.609 +/- 0.061	0.259 +/- 0.091	1.335 +/- 0.698	0.341 +/- 0.116
DR-CFR	0.657 +/- 0.028	0.240 +/- 0.032	0.789 +/- 0.091	0.261 +/- 0.036
DeR-CFR	**0.444 +/- 0.020**	**0.130 +/- 0.020**	**0.529 +/- 0.068**	**0.147 +/- 0.022**

Table 2: Ablation studies of DeR-CFR.

$\mathcal{L}_A$	$\mathcal{L}_I$	$\mathcal{L}_{C\_B}$	$\mathcal{L}_O$	PEHE	
				Within-sample	Out-of-sample
✓	✓	✓	✓	**0.444 +/- 0.020**	**0.529 +/- 0.068**
✓	✓	✓		0.478 +/- 0.033	0.542 +/- 0.053
✓	✓		✓	0.482 +/- 0.039	0.565 +/- 0.075
✓		✓	✓	0.479 +/- 0.030	0.560 +/- 0.071
	✓	✓	✓	0.635 +/- 0.035	0.858 +/- 0.133

Wu A, Yuan J, Kuang K, et al. Learning decomposed representations for treatment effect estimation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

Learning Decomposed Representation for Counterfactual Inference

Table 1: The results on IHDP.

IHDP				
Mean +/- Std	Within-sample		Out-of-sample	
Methods	PEHE	ϵ_{ATE}	PEHE	ϵ_{ATE}
CFR-MMD	0.702 +/- 0.037	0.284 +/- 0.036	0.795 +/- 0.078	0.309 +/- 0.039
CFR-WASS	0.702 +/- 0.034	0.306 +/- 0.040	0.798 +/- 0.088	0.325 +/- 0.045
CFR-ISW	0.598 +/- 0.028	0.210 +/- 0.028	0.715 +/- 0.102	0.218 +/- 0.031
SITE	0.609 +/- 0.061	0.259 +/- 0.091	1.335 +/- 0.698	0.341 +/- 0.116
DR-CFR	0.657 +/- 0.028	0.240 +/- 0.032	0.789 +/- 0.091	0.261 +/- 0.036
DeR-CFR	0.[illegible]	0.[illegible]	0.529 +/- 0.[illegible]	0.[illegible]

Table 2: Ablation studies of DeR-CFR.

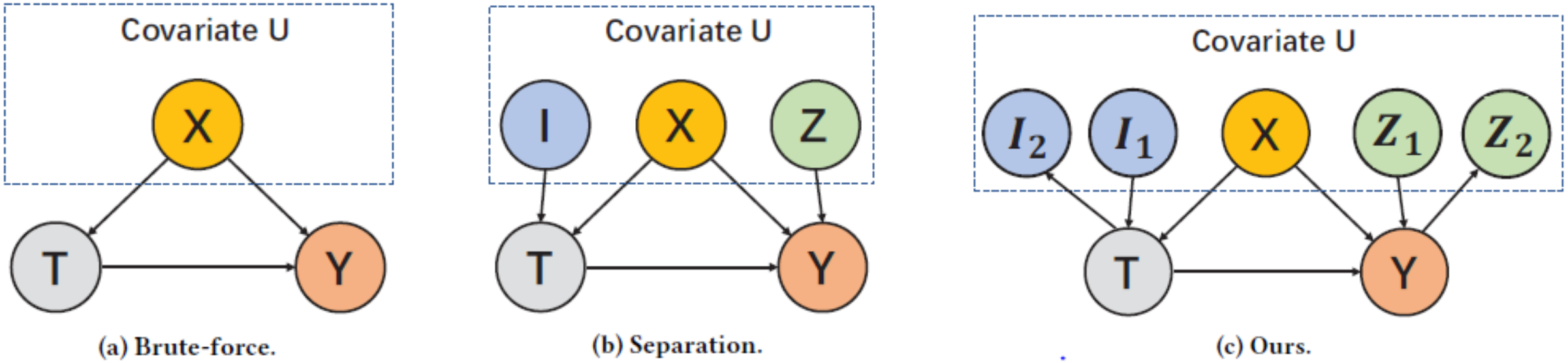
$\mathcal{L}_A$	$\mathcal{L}_I$	$\mathcal{L}_{C\_B}$	$\mathcal{L}_O$	PEHE	
				Within-sample	Out-of-sample
✓	✓	✓	✓	**0.444 +/- 0.020**	**0.529 +/- 0.068**
✓	✓	✓		0.478 +/- 0.033	0.542 +/- 0.053
✓	✓		✓	0.482 +/- 0.039	0.565 +/- 0.075
✓		✓	✓	0.479 +/- 0.030	0.560 +/- 0.071
	✓	✓	✓	0.635 +/- 0.035	0.858 +/- 0.133

Designed for pre-treatment/outcome variables.

How about with post-treatment/outcome variables?

Wu A, Yuan J, Kuang K, et al. Learning decomposed representations for treatment effect estimation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022.

Adjustment Feature Selection



(a) Brute-force. (b) Separation. (c) Ours.

Additional Assumption: NO mediator

Theoretical Results: Optimal features are the confounders and outcome-related covariates

How to select the optimal features?

Haotian Wang, Kun Kuang, et.al. Treatment Effect Estimation with Adjustment Feature Selection, KDD, 2023.

Adjustment Feature Selection

- Semi-parametric Inference
  - Definition: The statistical model M is indexed by
    - Parameter of Interest $\gamma(P)$: finite dimension (e.g., ATE)
    - Nuisance parameter: infinite dimension
    - Asymptotic Linear: $\sqrt{n}\big(T(n)-\gamma(P)\big)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}D_i(V)+o_p(1)$
      - ✓ Influence Function: $D(V)$
      - ✓ Estimator: $T(n)$ (e.g., Estimator of ATE)
    - Efficient Influence Function $D^{eff}(V)$ should achieve the Cramer-Rao Lower Bound (CRLB)
  - Efficient Influence Function of ATE estimation:
    - $D^{eff}(V)=\frac{\mathrm{I(T=1)-I(T=0)}}{\pi^{\mathrm{T}}(\mathrm{V})}\big(\mathrm{Y}-m_{\mathrm{V}}^{\mathrm{T}}(\mathrm{Y})\big)+m_{\mathrm{V}}^{\mathrm{T=1}}(\mathrm{Y})-m_{\mathrm{V}}^{\mathrm{T=0}}(\mathrm{Y})-\gamma(P)$
    - Where $\pi^{\mathrm{T}}(\mathrm{V})$: propensity score, $m_{\mathrm{V}}^{\mathrm{T}}(\mathrm{Y})$: regression model

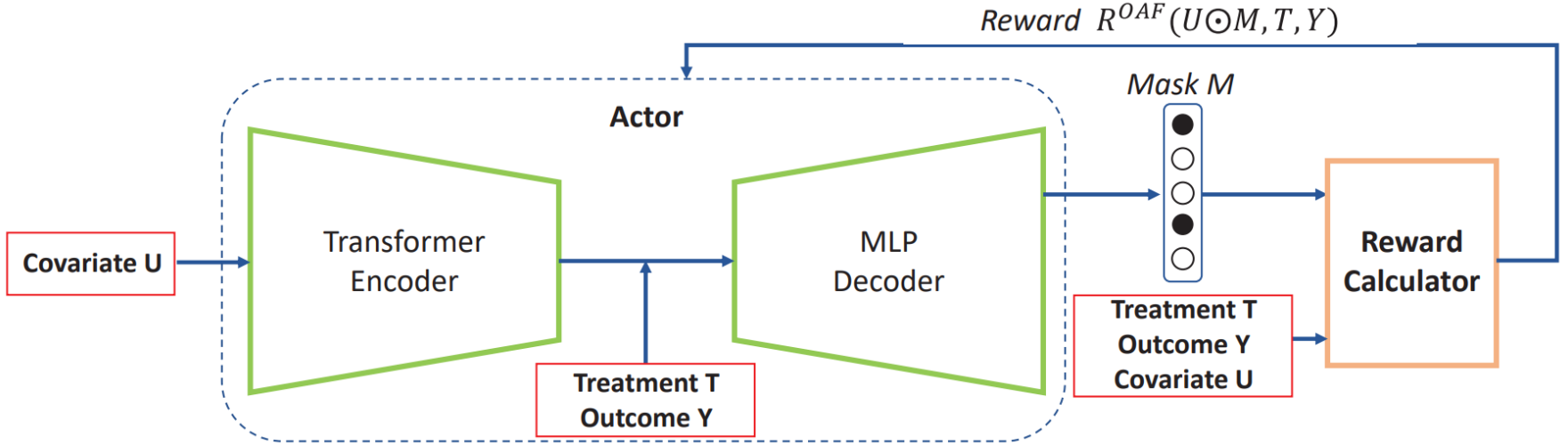
Adjustment Feature Selection

- Asymptotic Normality:
  - ✓ $E\left[D^{eff}(V)\right] = 0$
  - ✓ $\sqrt{n}(\gamma(\hat{P}) - \gamma(P)) \xrightarrow{d} N\left(0, \mathrm{Var}\left[D^{\mathrm{eff}}(\mathrm{V})\right]\right)$
- Our Goal: Identify the adjustment sets V that achieves minimal asymptotic variance $Var[D^{\mathrm{eff}}(\mathrm{V})]$
- Optimality of V: $Var[D^{\mathrm{eff}}(\mathrm{V})]$ is minimized if and only if V = {X, Z}
- Objective Function: $R^{\mathrm{OAF}}(V) = Var[D^{\mathrm{eff}}(\mathrm{V})]$
- Empirical Estimator: $\widehat{\mathrm{R}}^{\mathrm{OAF}}(\mathrm{V}) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{I(\mathrm{t}_i=1)-I(\mathrm{t}_i=0)}{\pi^{\mathrm{t}_i}(\mathrm{v}_i)}\left(\mathrm{y}_i - \overline{m^{\mathrm{t}_i}(\mathrm{v}_i)}\right)\right)^2$

Haotian Wang, Kun Kuang, et.al. Treatment Effect Estimation with Adjustment Feature Selection, KDD, 2023.

Adjustment Feature Selection

- How to select the optimal features V={X,Z}? It's a combinatorial optimization problem !
- Reinforcement Learning for Optimization:
  - Taking the variable selection as an end-to-end differentiable process
  - Using masks to select variables



Haotian Wang, Kun Kuang, et.al. Treatment Effect Estimation with Adjustment Feature Selection, KDD, 2023.

Adjustment Feature Selection

- Datasets: Linear Synthetic, Non-linear Synthetic, IHDP, Twins
- Metric: ☐ Feature selection accuracy: $Acc = \frac{|\widehat{M} - \mathbf{M}_0|_1}{d}$ ☐ MAE error: $\epsilon_{\text{ATE}} = |\text{ATE} - \widehat{\text{ATE}}|$

Dataset	Fs_Acc	R_err
S-20-l	95.0%	0.02
S-40-l	92.5%	0.04
S-80-l	90.0%	0.11
S-20-n	95.0%	0.06
S-40-n	95.0%	0.10
S-80-n	90.0%	0.13
IHDP	92.0%	0.11
Twins	94.7%	0.13

Settings		In_sample Prediction			Out_of_sample Prediction		
Feature Dimension		20	40	80	20	40	80
Statistical	Direct	4.69±0.62	7.09±0.68	8.92±0.76	5.23±0.41	6.28±1.41	9.28±1.32
	IPW	0.99±4.50	1.27±3.13	4.36±2.37	1.33±1.92	2.22±5.39	4.51±3.33
	AIPW	1.32±1.95	0.99±0.27	2.35±0.83	0.21±1.19	0.55±0.47	3.88±1.23
	TMLE	0.42±0.11	0.59±0.07	0.62±0.02	0.50±0.12	0.66±0.18	0.81±0.20
Machine	DragonNet	0.19±0.19	0.20±0.14	0.57±0.38	0.99±0.16	0.84±0.70	0.87±1.02
	GANITE	0.80±0.01	0.87±0.01	0.99±0.01	0.99±0.01	1.08±0.01	1.10±0.01
	DNOUT	0.47±0.01	0.62±0.04	0.92±0.09	0.50±0.02	0.61±0.05	0.95±0.09
	BART	0.92±0.20	2.03±0.27	2.89±0.98	0.92±0.20	2.25±0.16	2.98±1.10
Decomposed	AIPW_L	0.59±0.10	0.66±0.05	0.89±0.10	0.54±0.29	0.74±0.13	0.96±0.22
	DVD	0.95±0.03	0.83±0.01	0.76±0.01	1.06±0.08	0.64±0.01	1.05±0.73
	DR-CFR	0.88±0.08	1.18±0.16	2.08±0.69	1.28±0.08	1.69±0.73	1.52±0.51
	TEDVAE	0.37±0.01	0.43±0.02	0.55±0.03	0.38±0.03	0.49±0.04	0.60±0.02
	NICE	1.08±0.32	1.24±0.60	1.81±0.22	1.10±0.37	1.23±0.41	1.93±0.35
Ours	**OAFP_L**	**0.03±0.13**	**0.12±0.10**	**0.23±0.13**	**0.24±0.22**	**0.20±0.13**	**0.32±0.34**
	OAFP_N	**0.01±0.10**	**0.09±0.07**	**0.13±0.11**	**0.15±0.09**	**0.16±0.07**	**0.14±0.08**

Haotian Wang, Kun Kuang, et.al. Treatment Effect Estimation with Adjustment Feature Selection, KDD, 2023.

Challenge 2: Unobserved confounders

☐Standard Assumptions for Causal Inference:

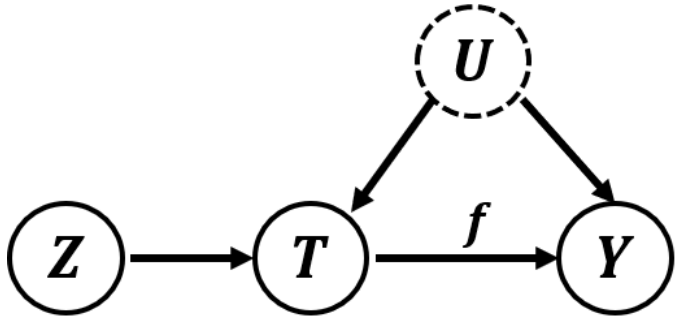
- **A2: Unconfounderness:** The distribution of treatment is independent of potential outcome when given the observed variables

$$T \perp \big(Y(0), Y(1)\big) | \, X$$

- In complex environments, NOT all confounder can be observed, i.e., the unconfounderness assumption is not satisfied.

- How to remove the bias from those unobserved confounders?

Instrumental Variable Regression



Conditions of IV (instrumental variable)

- Relevance: $P(T|Z) \neq P(T)$
- Exclusion: $P(Y|Z,T,U) \neq P(Y|T,U)$
- Unconfounded: $Z \perp U$

2SLS:

Stage 1: regressing T on Z $\quad \hat{T} = \hat{g}(Z)$

Stage 2: regressing Y on $\hat{T}$ $\quad \hat{Y} = \hat{f}(\hat{T})$

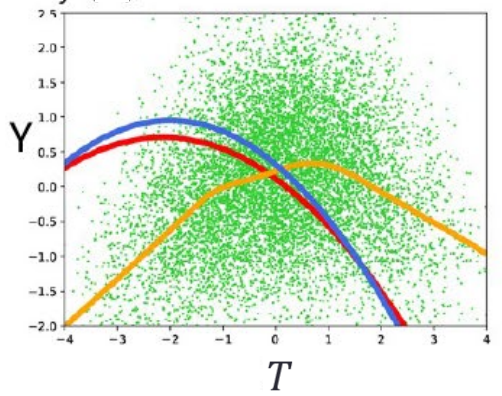
$Z \sim \mathcal{N}(0,1)$

$U \sim \mathcal{N}(0,1)$

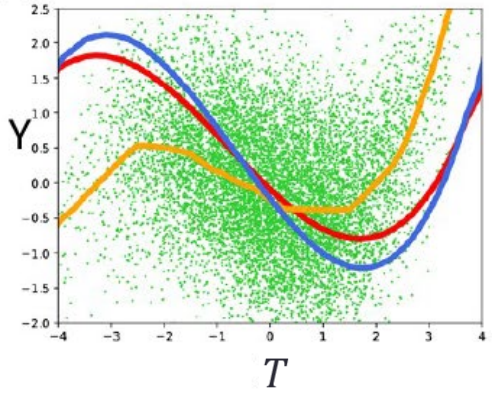
$T = Z + U$

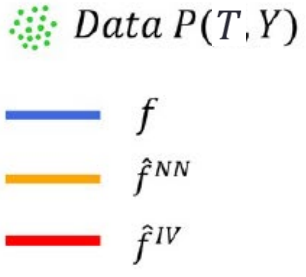
$Y = f(T) + U$

$f(T) = -0.1 \cdot T^2 - 0.4 \cdot T$



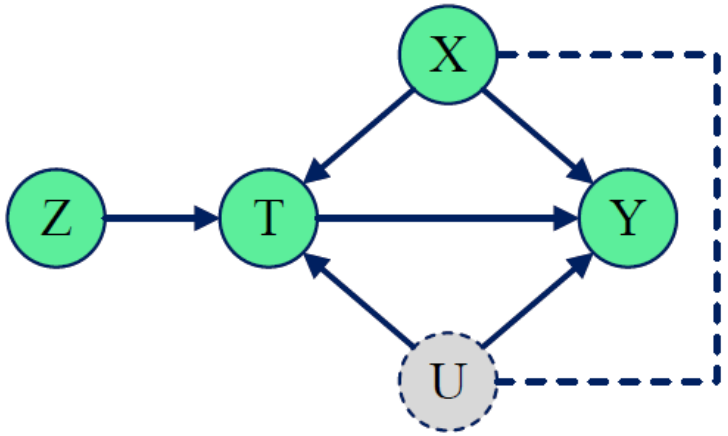
$f(T) = 0.05 \cdot T^3 + 0.1 \cdot T^2 - 0.8 \cdot T$

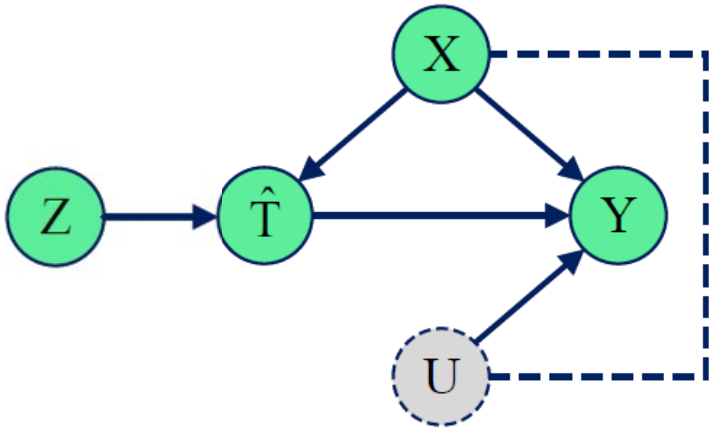




Requiring pre-defined IVs,
Limited to linear setting

Non-linear Instrumental Variable Regression





Non-linear IV regression (DeepIV, KernelIV et.al)

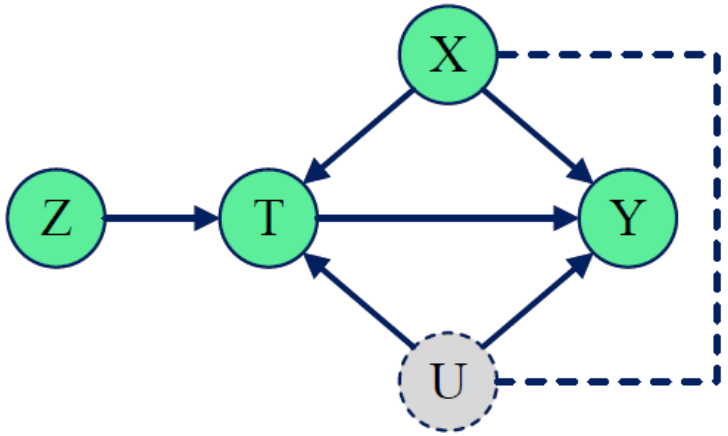
Stage 1: regressing T on Z and X $\quad \hat{T} = \hat{g}(Z, X)$

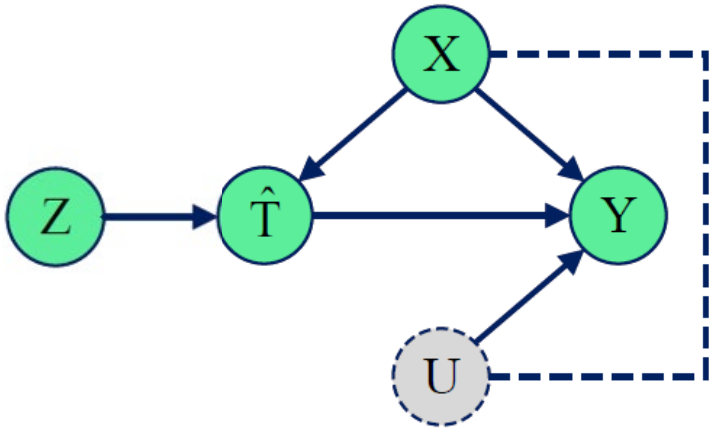
Stage 2: regressing Y on $\hat{T}$ and X $\quad \hat{Y} = \hat{f}(\hat{T}, X)$

Stage 1 regression brings confounding bias in stage 2

Confounder Balancing + IV Regression

Confounder Balanced Instrumental Variable Regression





CB-IV (Confounder Balanced IV regression):

Stage1 (Treatment regression): regressing T on Z and X $\quad \hat{T} = \hat{g}(Z, X)$

Confounder balancing: learning a balanced confounder representation $\phi(X)$ such that $\hat{T} \perp \phi(X)$

Stage 2 (Outcome regression): regressing Y on $\hat{T}$ and $\phi(X)$ $\quad \hat{Y} = \hat{f}(\hat{T}, \phi(X))$

Wu A, Kuang K, Li B, et al. Instrumental Variable Regression with Confounder Balancing, ICML 2022

Confounder Balanced Instrumental Variable Regression

Table 2: The bias (mean ± std) of ATE estimation on real-world data (Data-m_Z-m_X-m_U)

Within-Sample				
Method	**IHDP-2-6-0**	**IHDP-2-4-2**	**Twins-5-8-0**	**Twins-5-5-3**
DeepIV-LOG	2.8736 ± 0.0577	2.6227 ± 0.0651	0.0135 ± 0.0215	0.0237 ± 0.0111
DeepIV-GMM	3.7760 ± 0.0316	3.7396 ± 0.0402	0.0194 ± 0.0047	0.0221 ± 0.0041
OneSIV	1.7249 ± 0.3752	1.7411 ± 0.3422	0.0083 ± 0.0191	0.0080 ± 0.0167
DFIV	3.5543 ± 0.0891	3.6218 ± 0.1038	0.0268 ± 0.0005	0.0265 ± 0.0003
DFL	3.2018 ± 0.0496	3.1991 ± 0.0374	0.0624 ± 0.0586	0.0847 ± 0.0049
DirectRep	0.0675 ± 0.0562	0.4600 ± 0.0711	0.0167 ± 0.0171	0.0193 ± 0.0251
CFR	0.0854 ± 0.0579	0.4826 ± 0.0642	0.0115 ± 0.0167	0.0223 ± 0.0176
DRCFR	0.0553 ± 0.0644	0.4336 ± 0.0692	0.0114 ± 0.0221	0.0118 ± 0.0174
CB-IV	**0.0117 ± 0.3882**	**0.1601 ± 0.2499**	**0.0067 ± 0.0271**	**0.0014 ± 0.0249**
Out-of-Sample				
Method	**IHDP-2-6-0**	**IHDP-2-4-2**	**Twins-5-8-0**	**Twins-5-5-3**
DeepIV-LOG	2.8760 ± 0.0553	2.6226 ± 0.0692	0.0140 ± 0.0208	0.0238 ± 0.0111
DeepIV-GMM	3.7768 ± 0.0350	3.7388 ± 0.0416	0.0193 ± 0.0047	0.0221 ± 0.0040
OneSIV	1.7287 ± 0.3725	1.7351 ± 0.3430	0.0082 ± 0.0191	0.0081 ± 0.0168
DFIV	3.5538 ± 0.0904	3.6225 ± 0.1061	0.0268 ± 0.0005	0.0265 ± 0.0003
DFL	3.2038 ± 0.0496	3.1994 ± 0.0376	0.0624 ± 0.0584	0.0846 ± 0.0046
DirectRep	0.0608 ± 0.0817	0.4571 ± 0.0759	0.0162 ± 0.0175	0.0194 ± 0.0253
CFR	0.0785 ± 0.0810	0.4804 ± 0.0687	0.0110 ± 0.0163	0.0225 ± 0.0180
DRCFR	0.0450 ± 0.0953	0.4321 ± 0.0673	0.0113 ± 0.0219	0.0118 ± 0.0174
CB-IV	**0.0150 ± 0.3927**	**0.1578 ± 0.2540**	**0.0065 ± 0.0270**	**0.0015 ± 0.0247**

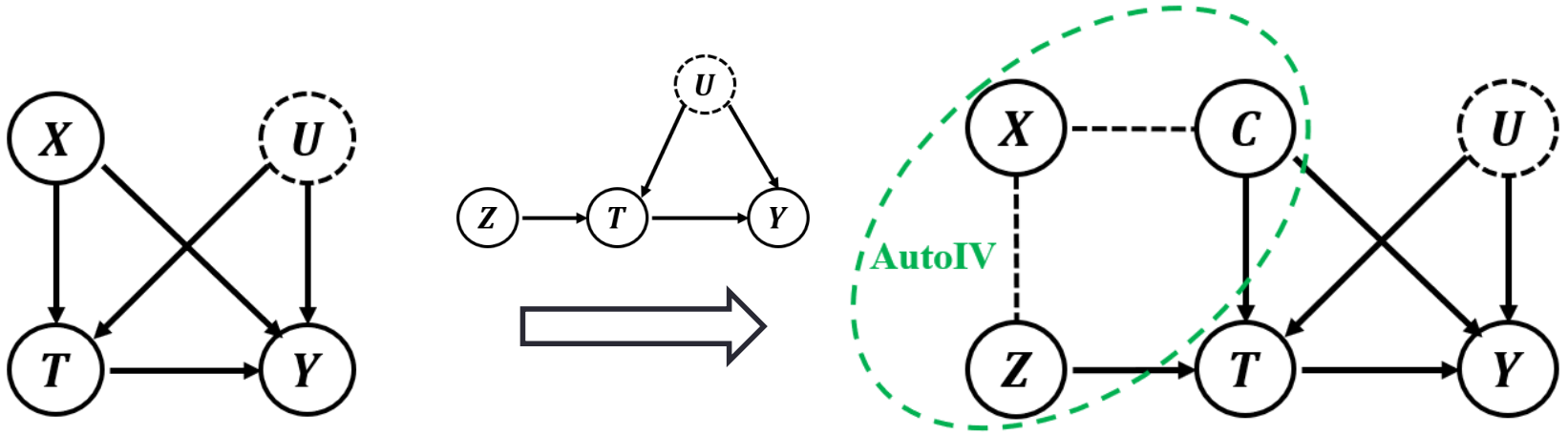
IV based methods

Confounder balancing based methods

Requiring pre-defined IVs

Wu A, Kuang K, Li B, et al. Instrumental Variable Regression with Confounder Balancing, ICML 2022

AutoIV: Counterfactual Learning with Unobserved Confounders via Automatically generating IVs



Conditions of IV

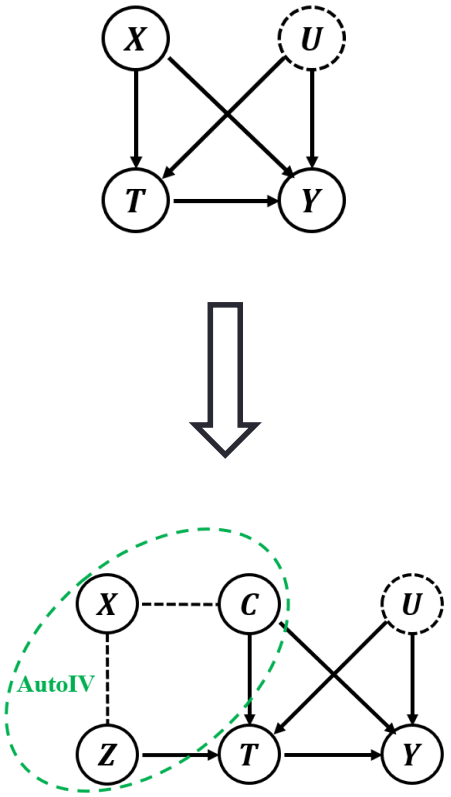
- Relevance: $P(T|Z) \neq P(T)$
- Exclusion: $P(Y|Z,T,C) \neq P(Y|T,C)$
- Unconfounded: $Z \perp C$

⇦ Mutual Information Representation Learning

But exclusion might not be satisfied

Yuan J, Wu A, Kuang K, et al. Auto IV: Counterfactual Prediction via Automatic Instrumental Variable Decomposition[J]. TKDD, 2022.

AutoIV: Counterfactual Learning with Unobserved Confounders via Automatically generating IVs



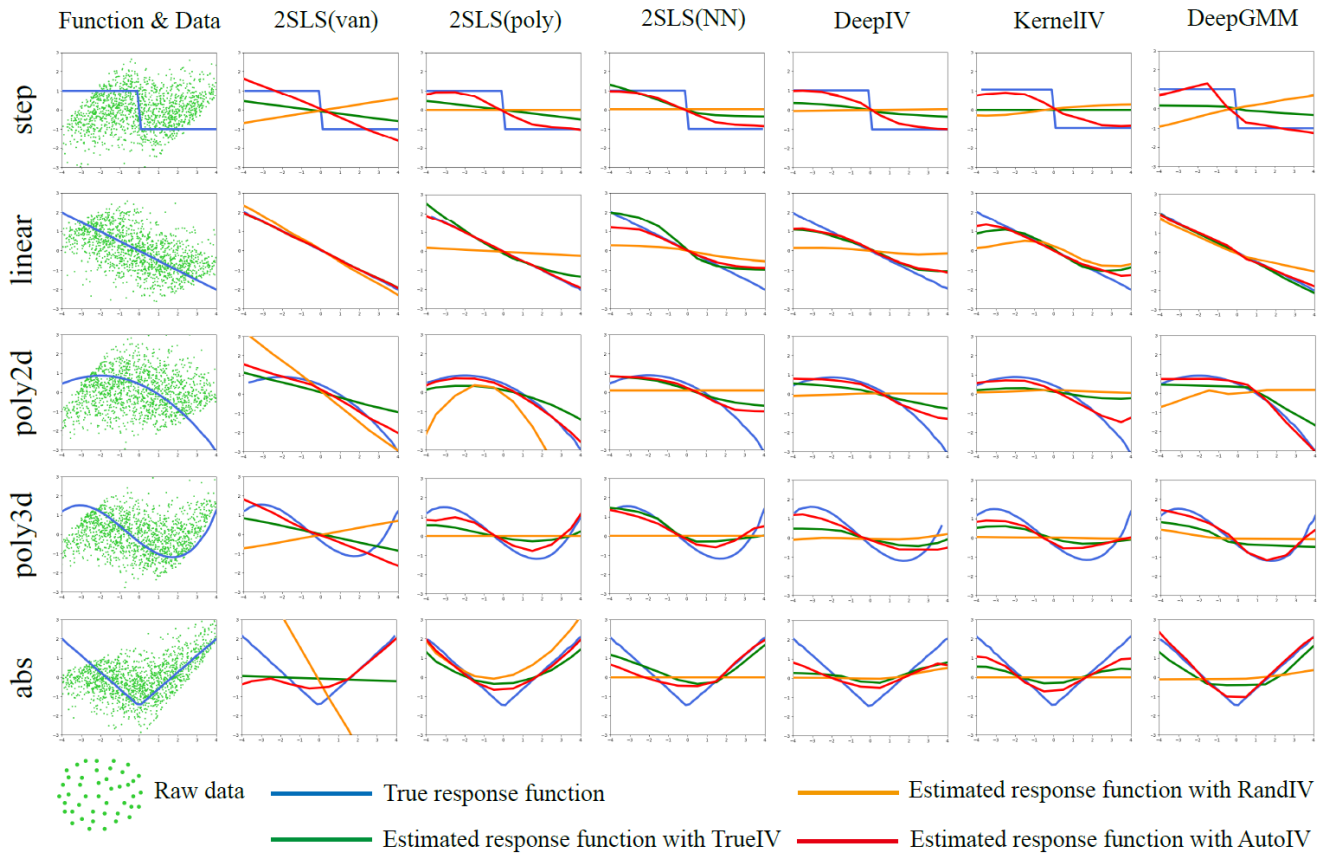


Figure 2: Response function prediction in low-dimensional scenarios.

Yuan J, Wu A, Kuang K, et al. Auto IV: Counterfactual Prediction via Automatic Instrumental Variable Decomposition[J]. TKDD, 2022.

Challenge 3: Complex Treatments

☐Standard Assumptions for Causal Inference:

- **A1: Stable Unit Treatment Value (SUTVA):** The effect of treatment on a unit is independent of the treatment assignment of other units

$$P(Y_i|T_i, T_j, X_i) = P(Y_i|T_i, X_i)$$

- In complex environments, for example, in social network, the treatment might not satisfy the **SUTVA** assumption.

- How to precisely estimate the effect of complex treatments?

NetIV: Networked Instrumental Variable for Treatment Effect Estimation with Unobserved Confounders

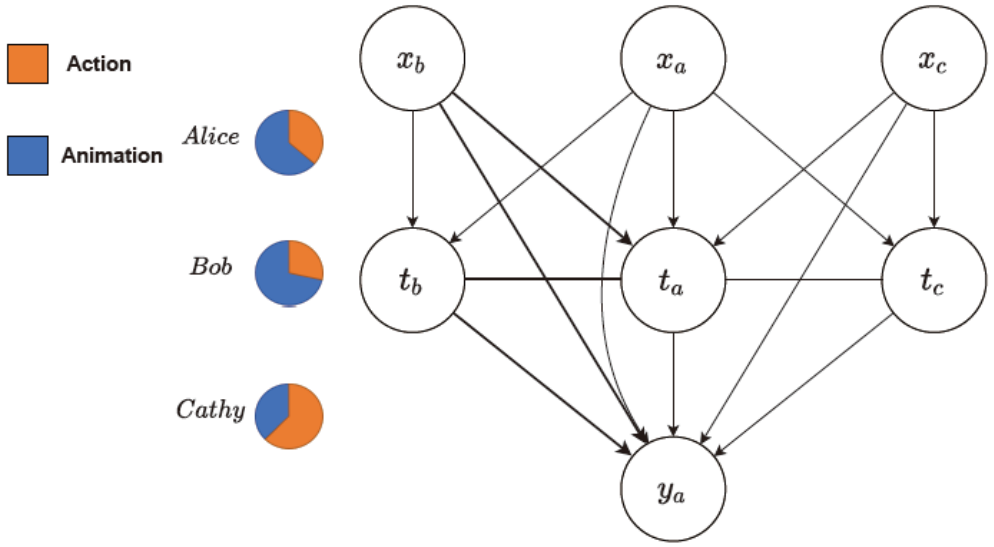


Figure 1: A motivating example to illustrate the setting of heterogeneous interference in networks.

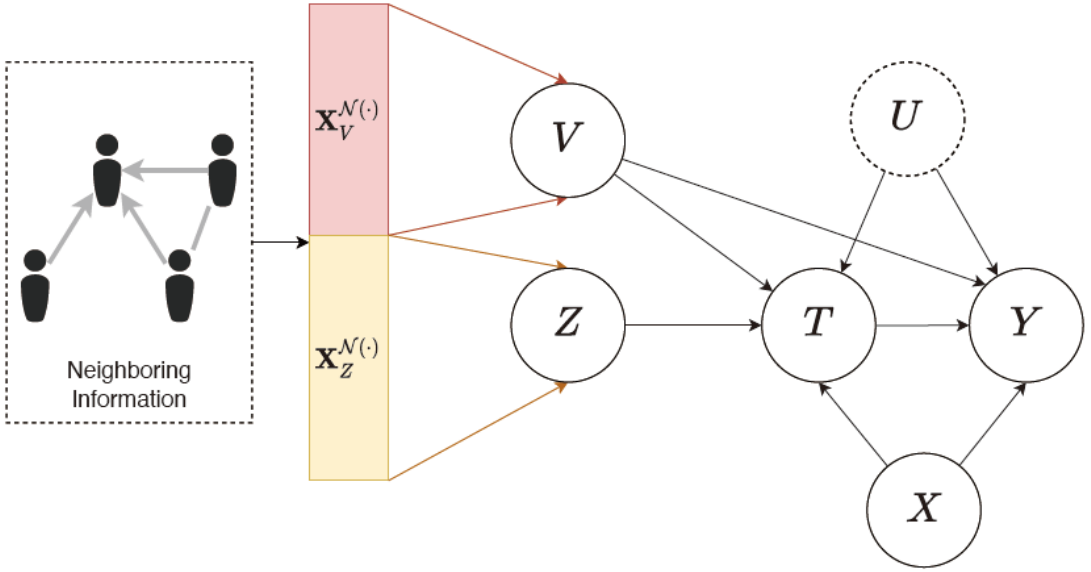
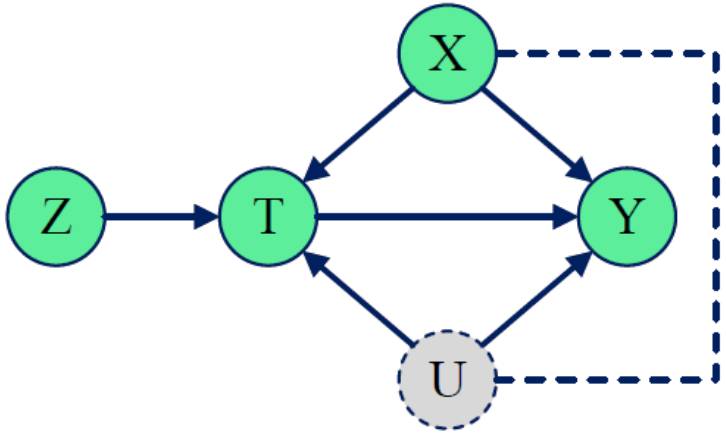


Figure 1: Causal diagram of the proposed NetIV framework. NetIV learns representation from neighboring information to serve as the role of IV Z and confounder proxy V.

A part of neighboring information can serve as the role of IV, called NetIV.

Ziyu Zhao, Kun Kuang, Fei Wu, et.al. NetIV: Networked Instrumental Variable for Treatment Effect Estimation with Unobserved Confounders, working paper.

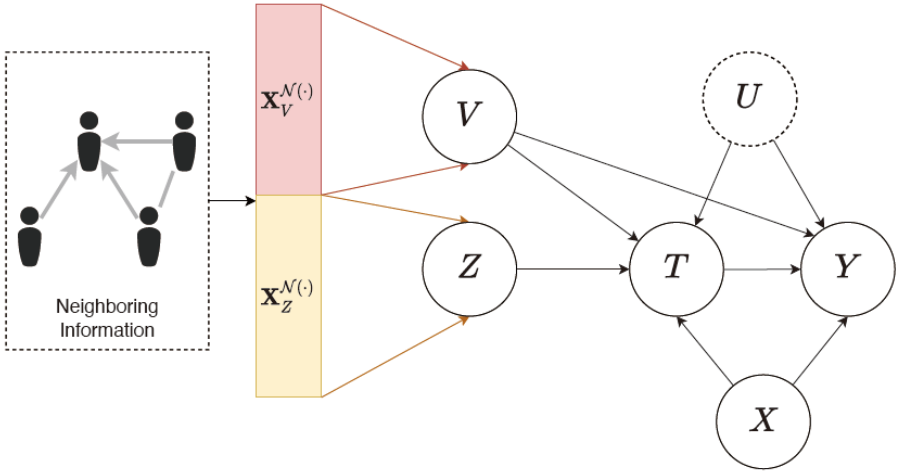
NetIV: Networked Instrumental Variable for Treatment Effect Estimation with Unobserved Confounders



Non-linear IV regression (DeepIV, KernelIV et.al)

Stage 1: regressing T on Z and X $\quad \hat{T} = \hat{g}(Z, X)$

Stage 2: regressing Y on $\hat{T}$ and X $\quad \hat{Y} = \hat{f}(\hat{T}, X)$



Proposed NetIV regression:

Stage 1: regressing T on $\{Z, V\}$ and X $\quad \hat{T} = \hat{g}(\{Z, V\}, X)$

Stage 2: regressing Y on $\hat{T}$ and X $\quad \hat{Y} = \hat{f}(\hat{T}, X)$

Stage 2 might bring bias of V if the model is mis-specified.

Ziyu Zhao, Kun Kuang, Fei Wu, et.al. NetIV: Instrumental Variable Regression for Treatment Effect Estimation with Networked Observational Data, working paper.

Summary: Causal Inference in Complex Environments

- Challenge 1: High dimensional variables, but NOT all variables are confounders.

✓DeR-CFR, OFA: confounders and adjustment features selection

- Challenge 2: Unobserved confounders, NOT all confounders are observed. IV based method is a great approach for the problem, but limited to linear and requires pre-defined IV.

✓CB-IV: from linear-IV regression to Non-linear IV regression

✓AutoIV: generating a representation to serve as the role of IV

- Challenge 3: Complex Treatments without SUTVA assumption

✓NetIV: a part of neighboring information to serve as the role of IV

IVs in Causal Inference and Machine Learning

Instrumental Variables in Causal Inference and Machine Learning: A Survey

Anpeng Wu, Kun Kuang, Ruoxuan Xiong, Fei Wu, *Senior Member, IEEE*

Abstract—Causal inference is the process of using assumptions, study designs, and estimation strategies to draw conclusions about the causal relationships between variables based on data. This allows researchers to better understand the underlying mechanisms at work in complex systems and make more informed decisions. In many settings, we may not fully observe all the confounders that affect both the treatment and outcome variables, complicating the estimation of causal effects. To address this problem, a growing literature in both causal inference and machine learning proposes to use Instrumental Variables (IV). This paper serves as the first effort to systematically and comprehensively introduce and discuss the IV methods and their applications in both causal inference and machine learning. First, we provide the formal definition of IVs and discuss the identification problem of IV regression methods under different assumptions. Second, we categorize the existing work on IV methods into three streams according to the focus on the proposed methods, including two-stage least squares with IVs, control function with IVs, and evaluation of IVs. For each stream, we present both the classical causal inference methods, and recent developments in the machine learning literature. Then, we introduce a variety of applications of IV methods in real-world scenarios and provide a summary of the available datasets and algorithms. Finally, we summarize the literature, discuss the open problems and suggest promising future research directions for IV methods and their applications. We also develop a toolkit of IVs methods reviewed in this survey at https://github.com/causal-machine-learning-lab/mliv.

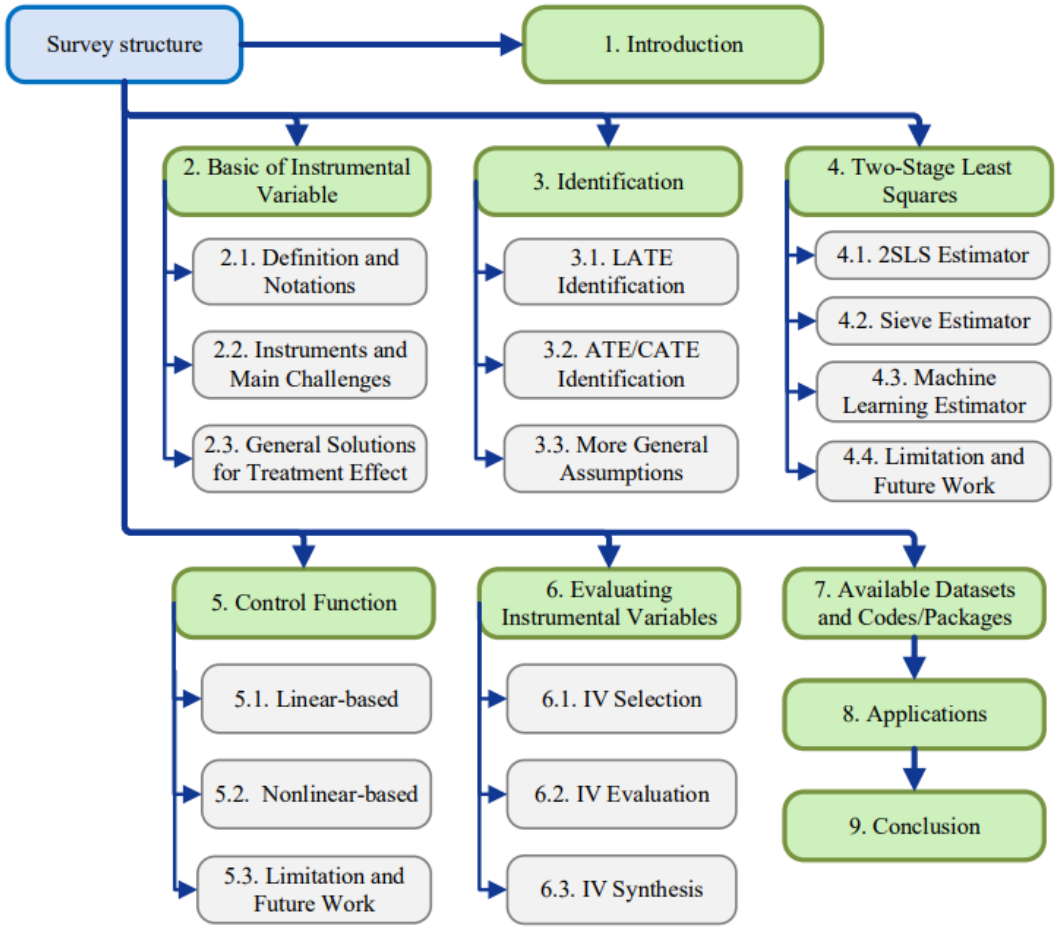


Fig. 2: Outline of the Survey.

Anpeng Wu, Kun Kuang, Ruoxuan Xiong, Fei Wu, Instrumental Variables in Causal Inference and Machine Learning: A Survey[J]. arXiv preprint arXiv:2212.05778, 2022.

IVs in Causal Inference and Machine Learning

mliv

```
from mliv.dataset.demand import gen_data
from mliv.utils import CausalDataset
gen_data()
data = CausalDataset('./Data/Demand/0.5_1.0_0.0_10000/1/')

from mliv.inference import Vanilla2SLS
from mliv.inference import Poly2SLS
from mliv.inference import NN2SLS
from mliv.inference import OneSIV
from mliv.inference import KernelIV
from mliv.inference import DualIV
from mliv.inference import DFL
from mliv.inference import AGMM
from mliv.inference import DeepGMM
from mliv.inference import DFIV
from mliv.inference import DeepIV           # Tensorflow & keras

for mod in [OneSIV,KernelIV,DualIV,DFL,AGMM,DeepGMM,DFIV,Vanilla2SLS,Poly2SLS,NN2SLS]:
    model = mod()
    model.config['num'] = 100
    model.config['epochs'] = 10
    model.fit(data)

    print(mod)
```

Anpeng Wu, Kun Kuang, Ruoxuan Xiong, Fei Wu, Instrumental Variables in Causal Inference and Machine Learning: A Survey[J]. arXiv preprint arXiv:2212.05778, 2022.

Thank You!

Kun Kuang
kunkuang@zju.edu.cn
Homepage: https://kunkuang.github.io/