Causal Recommendation: Progresses and Future Directions

Lecture Tutorial for SIGIR 2023

Organizers: Wenjie Wang, Yang Zhang, Haoxuan Li, Peng Wu, Fuli Feng, Xiangnan He

23 July 2023

Webpage: https://causalrec.github.io/

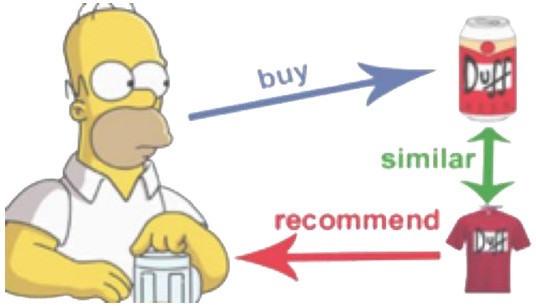
Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

Information Seeking

- **Information explosion era**
  - E-commerce: 12 million items in Amazon.
  - Social networks: 2.8 billion users in Facebook.
  - Content sharing platforms: 720,000 hours videos uploaded to Youtube per day.

- **Recommender system**



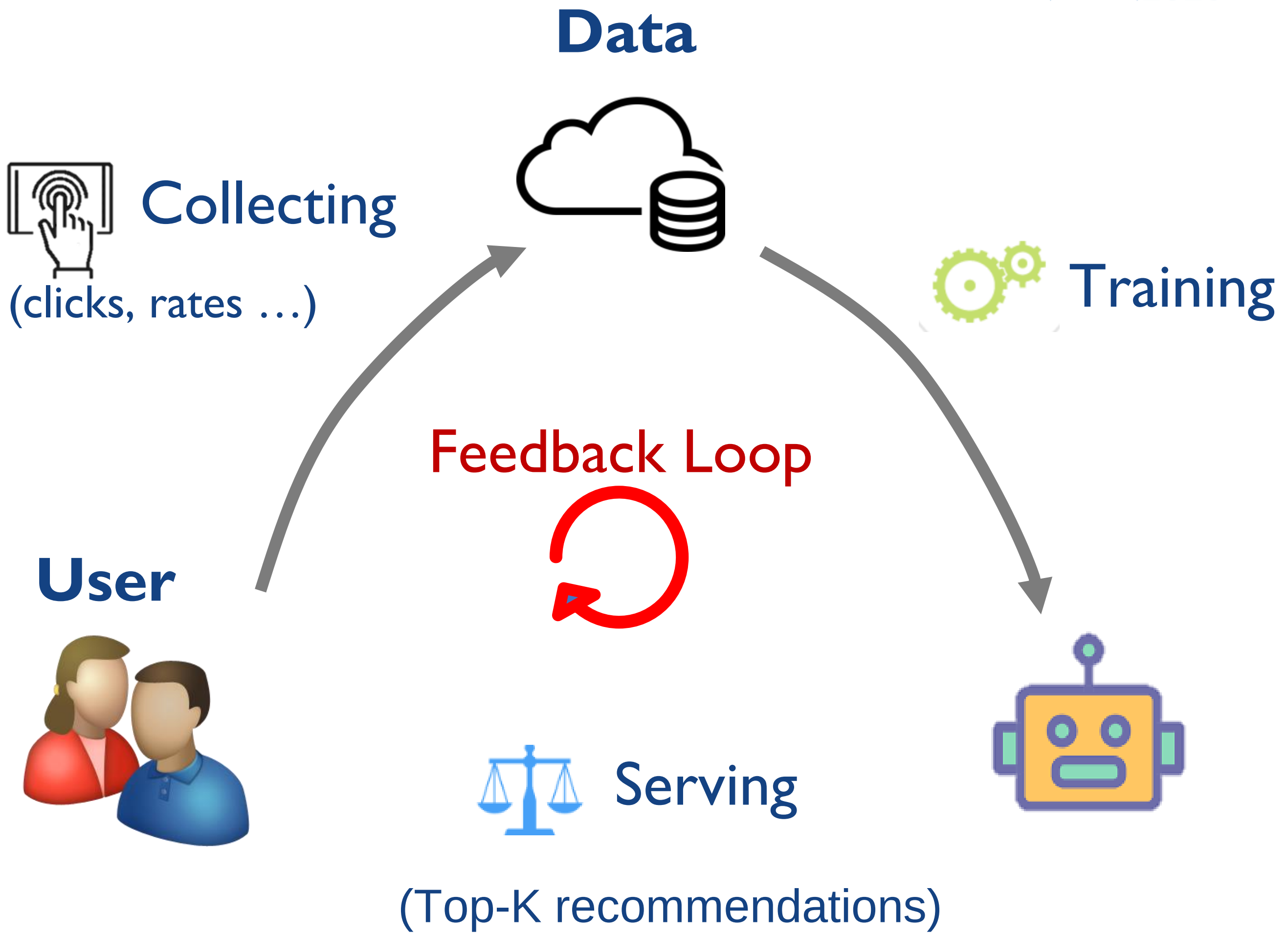
Recommendation

Information seeking
via **implicit feedback**

Recommender system is a powerful tool to address information overload.

Ecosystem of RecSys

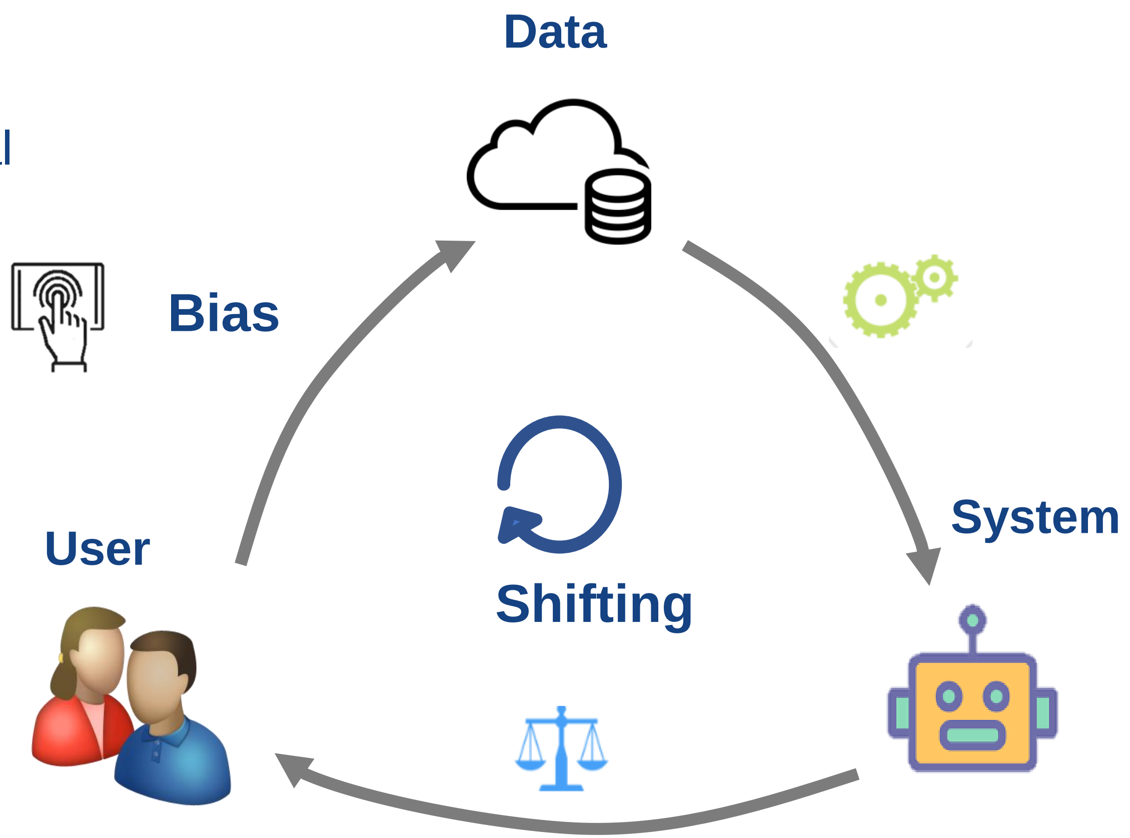
- Workflow of RecSys
  - **Training**: RecSys is trained on observed user-item interactions.
  - **Serving**: RecSys infers user preference over items and recommend Top-K items.
  - **Collecting**: collect user interactions on the recommended items for further training.
- Forming a feedback loop



Chen et al. arxiv 2021. Bias and Debias in Recommender System: A Survey and Future Directions

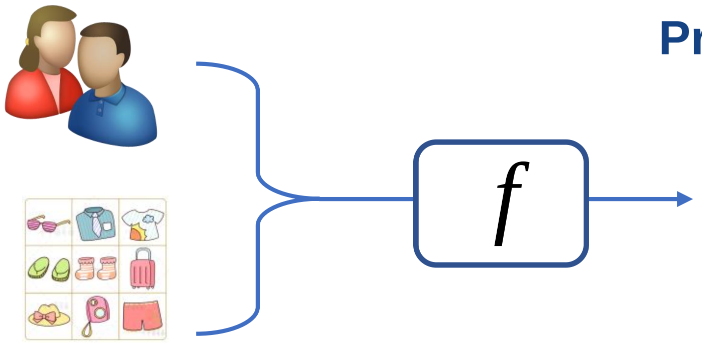
Shortcomings of Data-driven RecSys

- **Bias in data (collecting):**
  - Data is observational rather than experimental (missing-not-at-random)
  - Affected by many hidden factors:
    - Public opinions, *etc.*
- **Bias shifting along time:**
  - User/item feature changes
    - Income, marriage status
  - Preference shifting



Fitting Historical Data

- Minimizing the difference between historical feedback and model prediction



Predicted Score

3	2	1	4
2	2	3	5
2	4	5	5

$\approx$

History feedback

3		1	
2	3		5
2		5	

- **Collaborative filtering:** Similar users perform similarly in future

Shallow representation learning
- Matrix factorization & factorization machines

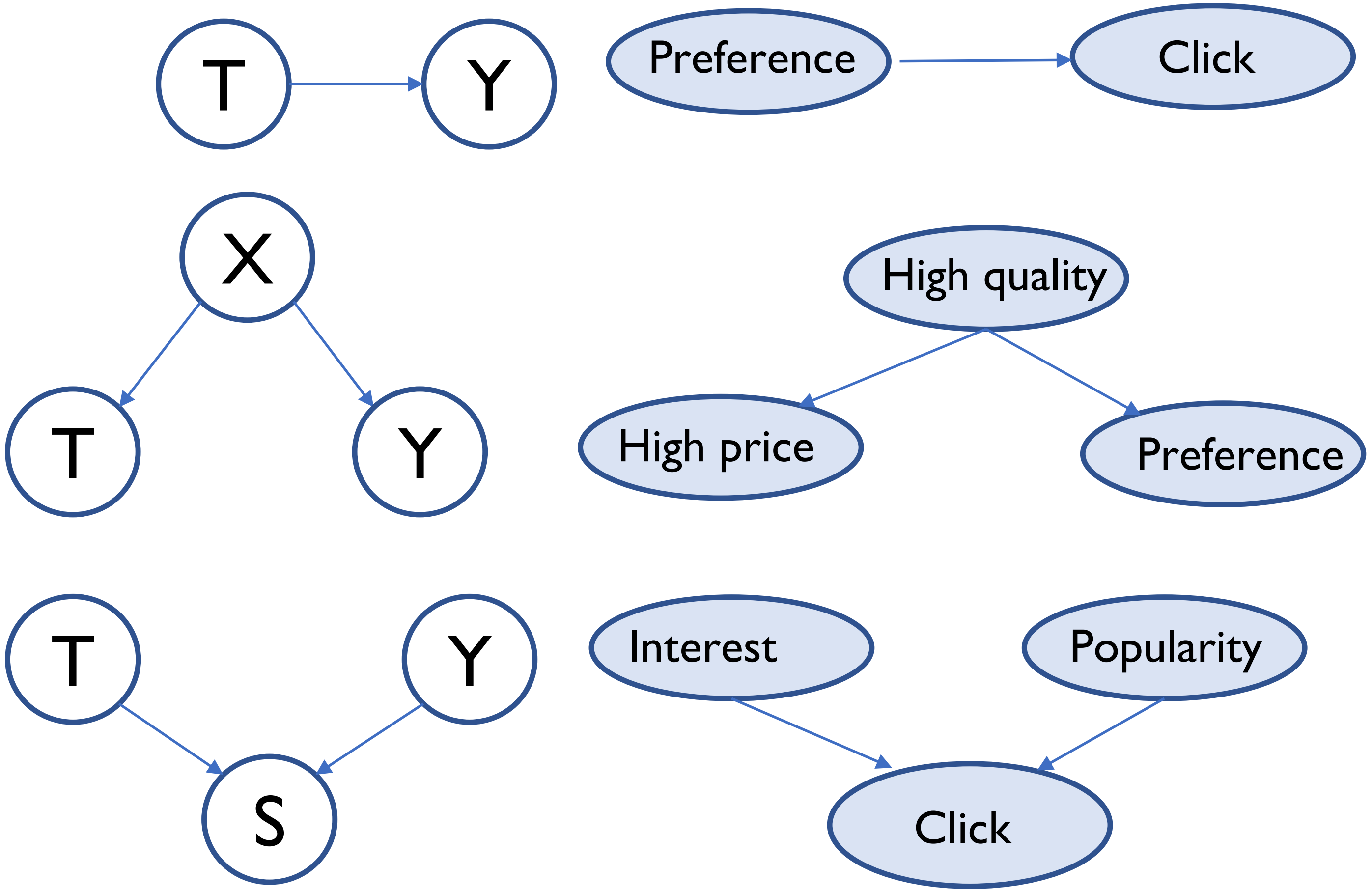
	Feature vector x																					Target y	
	User				Movie					Other Movies rated					Time	Last Movie rated							
	A	B	C	...	TI	NH	SW	ST	...	TI	NH	SW	ST	...		TI	NH	SW	ST	...			
$x^{(1)}$	1	0	0	...	1	0	0	0	...	0.3	0.3	0.3	0	...	13	0	0	0	0	...		5	$y^{(1)}$
$x^{(2)}$	1	0	0	...	0	1	0	0	...	0.3	0.3	0.3	0	...	14	1	0	0	0	...		3	$y^{(2)}$
$x^{(3)}$	1	0	0	...	0	0	1	0	...	0.3	0.3	0.3	0	...	16	0	1	0	0	...		1	$y^{(2)}$
$x^{(4)}$	0	1	0	...	0	0	1	0	...	0	0	0.5	0.5	...	5	0	0	0	0	...		4	$y^{(3)}$
$x^{(5)}$	0	1	0	...	0	0	0	1	...	0	0	0.5	0.5	...	8	0	0	1	0	...		5	$y^{(4)}$
$x^{(6)}$	0	0	1	...	1	0	0	0	...	0.5	0	0.5	0	...	9	0	0	0	0	...		1	$y^{(5)}$
$x^{(7)}$	0	0	1	...	0	0	1	0	...	0.5	0	0.5	0	...	12	1	0	0	0	...		5	$y^{(6)}$

Neural representation learning
- Neural collaborative filtering
- Graph neural networks
- Sequential model
- Textual & Visual encoders

Learning correlations between input features and interactions.

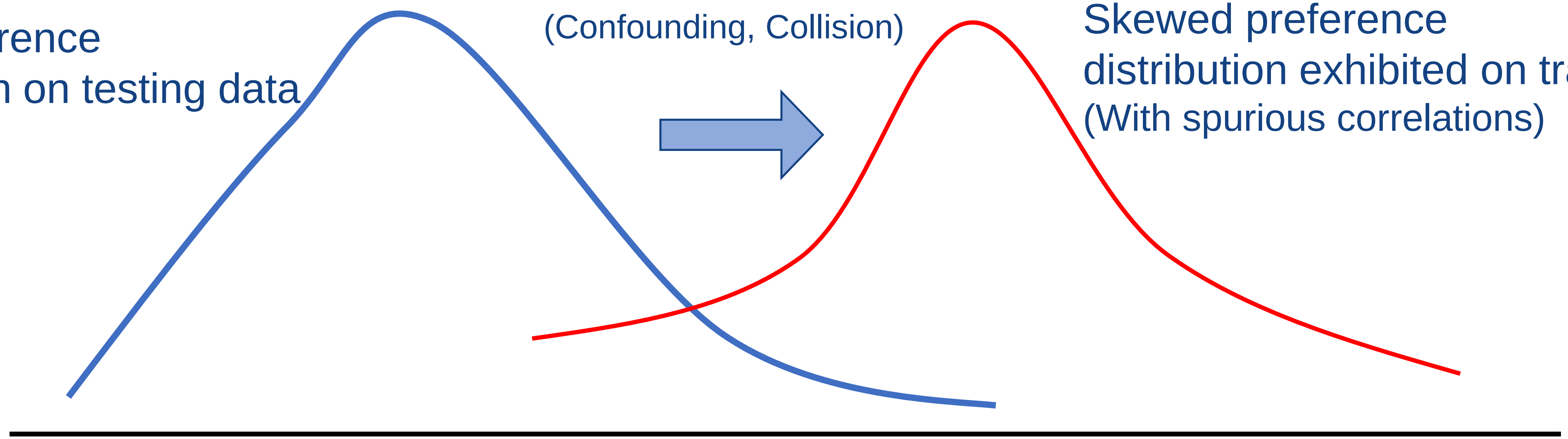
Shortcomings of Data-driven RecSys

- Correlation != preference: Correlations may not reflect the true causes of interactions.
- Three basic types of correlations:
  - **Causation**
    - Stable and explainable
  - **Confounding**
    - Ignoring X
    - Spurious correlation
  - **Collision**
    - Condition on S
    - Spurious correlation



Shortcomings of Data-driven RecSys

- Data-driven methods will learn skewed user preference:



- Data-driven methods may infer spurious correlations, which deviates from users' true preference.

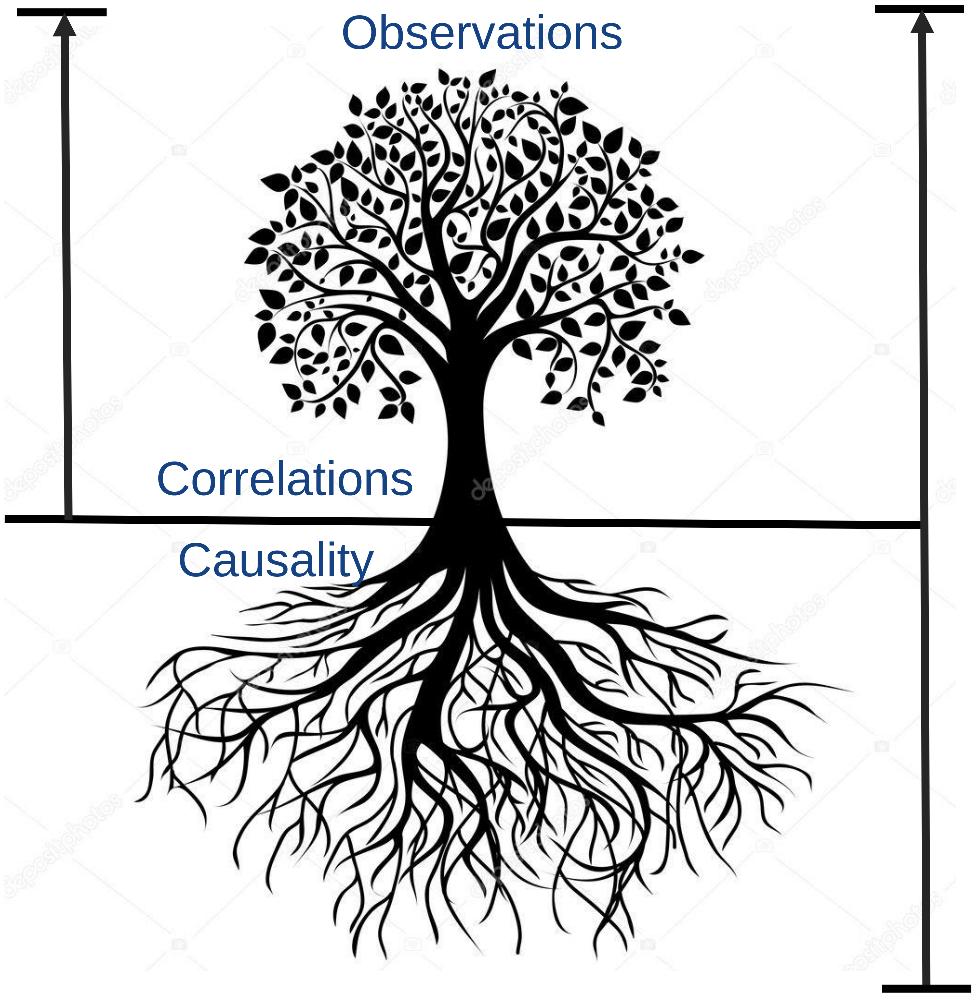
Correlation != preference

Chen et al. TOIS 2023. Bias and Debias in Recommender System: A Survey and Future Directions

Why Causal Inference?

- Aim: Understanding the inherent **causal mechanism** behind user behaviors
  - Capturing user true preference

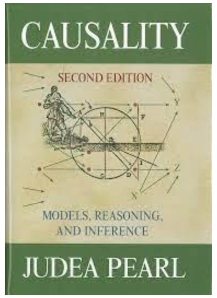
- Making reliable & explainable recommendations
  - Correlation + Causality > Correlation



Classification of Causal Recommendation

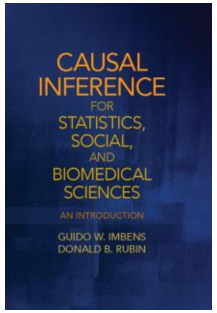
- Structural Causal Model (SCM)

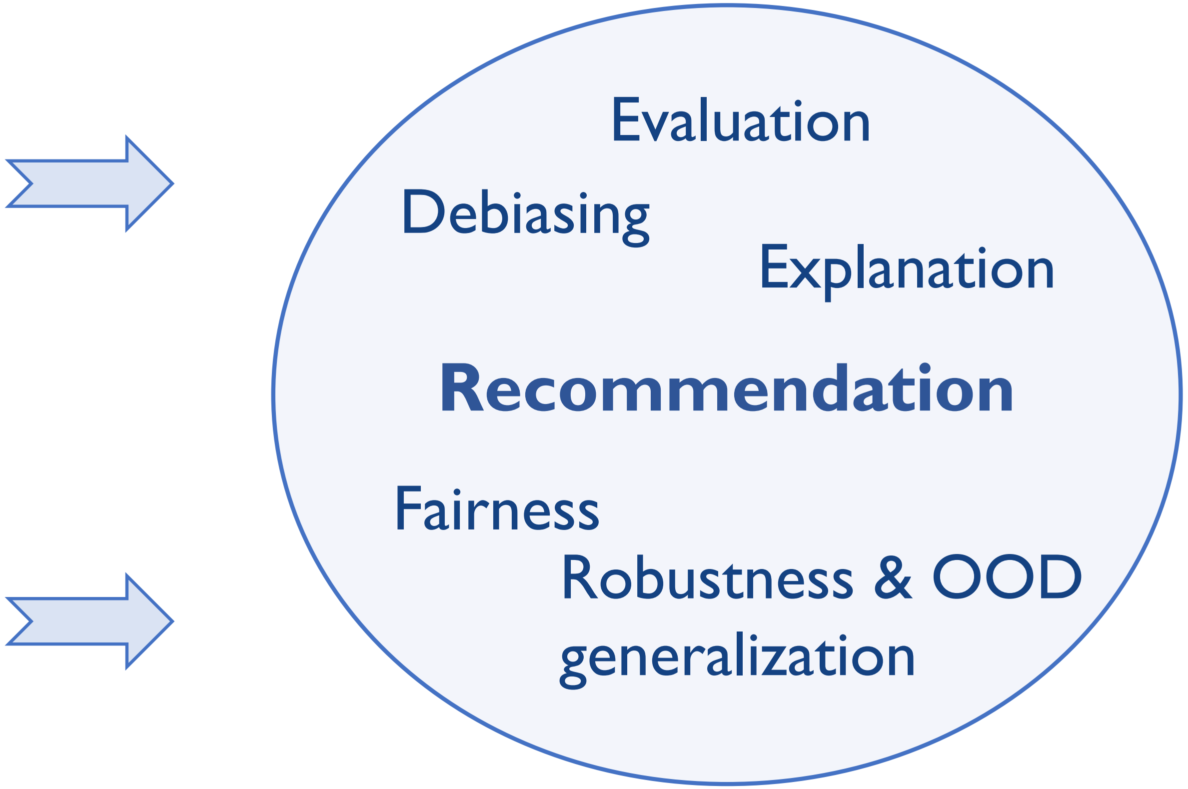
(Judea Pearl)



- Potential Outcome Framework

(Donald B. Rubin)





Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

Structural Causal Model

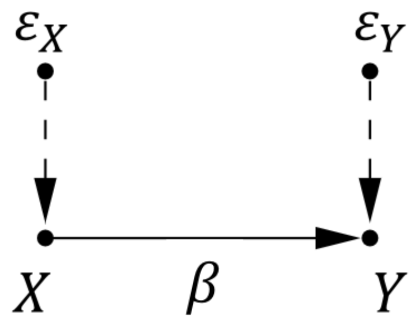
- How can common understandings, such as the fact that symptoms do not cause diseases, be expressed mathematically?

X: disease   *Y: symptom*

$$X = U_X$$
$$Y = \beta X + U_Y$$

U_X *and* U_Y: exogenous

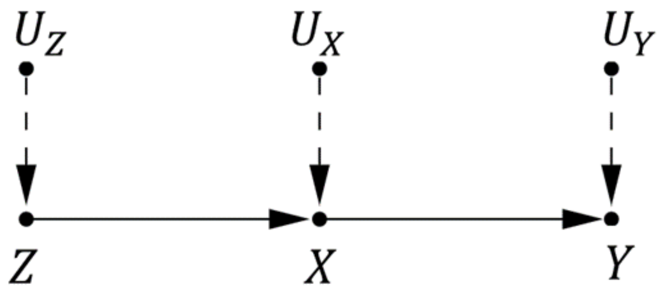
To express the inherent directionality



Causal Graph / Causal Diagram

Causal diagrams encodes causal assumption via missing arrows, representing claims of zero influences

- General form:



Non-parametric interpretation

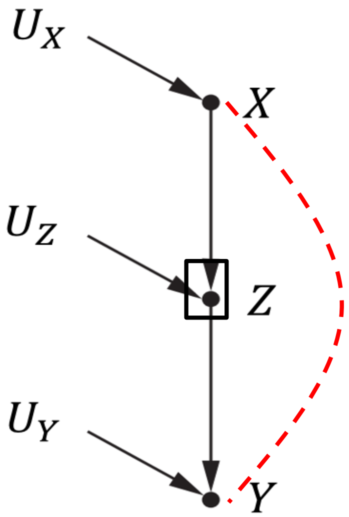
$$Z = f_Z(U_Z)$$
$$X = f_X(Z, U_X)$$
$$Y = f_Y(X, U_Y)$$

Pearl, Judea. "Causal inference." Causality: objectives and assessment (2010): 39-58.

Structural Causal Model

- Basic causal structures in causal graph

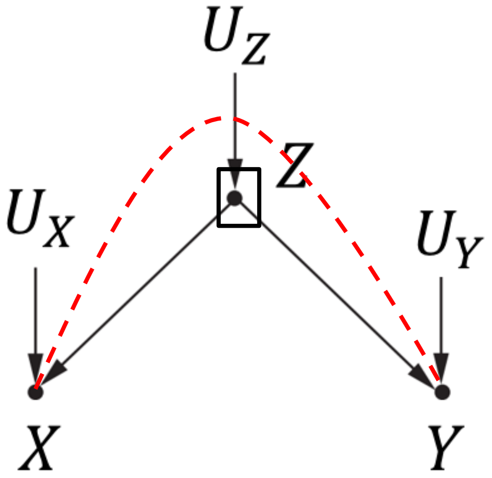
Chain



Z: **mediator**

- X and Y are associated.
- condition on Z, X and Y are independent.

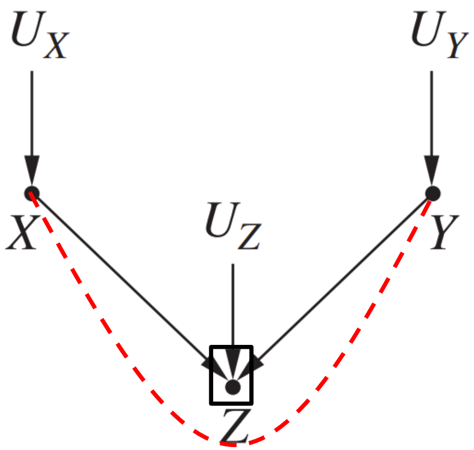
Confounding



Z: **confounder**

- X does not affect Y, but X and Y are correlated. (Spurious correlations).
- condition on Z, X and Y are independent, blocking the spurious correlations.

Colliding



Z: **collider**

- X and Y are independent.
- Condition on Z, X and Y are correlated, bringing spurious correlations.

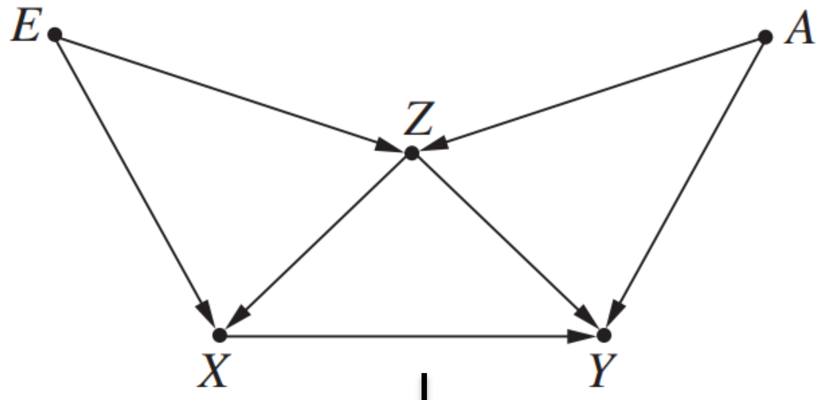
Structural Causal Model

- Correlation is not causation

Confounders and controlling colliders would bring spurious correlations between treatment and outcome.

It is impossible to answer causal question with correlation-level tools

Confounder E,Z,A will bring spurious correlations



$$P(Y|do(X)) = \sum_{z,a} P(Y|X,z,a)P(z,a)$$

- *do*-calculus

It provides various principles to identify target causal effect.
For example, utilize *the backdoor adjustment when confounders exist*

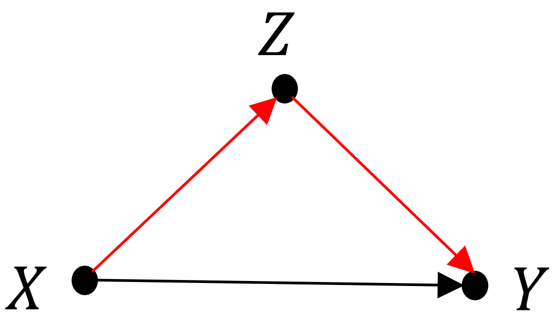
If any node in Z isn't a descendant of X, and Z blocks every path between X and Y that contains an arrow into X (backdoor path), then the average causal effect of X on Y is:

$$P(Y|do(X)) = \sum_Z P(Y|X,Z)P(Z)$$

Structural Causal Model

- SCM provides both a mathematical foundation and a friendly calculus for the analysis of causal effects and counterfactuals.
- It can deal with the estimation of three types of causal queries:

  - Queries about the effect of potential interventions.
    To compute causal effect, e.g., $P(Y|do(X))$
  - Queries about counterfactuals.
    e.g., whether event A would occur **if event B had been different**?
  - Queries about the direct / indirect effects. (based on counterfactuals)



the direct effects of X on Y: $X \to Y$
the indirect effects of X on Y: $X \to Z \to Y$

SCM for Recommendation

Q1: Queries about causal effect.

Q2: Queries about counterfactuals.

Q3: Queries about the direct/indirect effects.

Causal queries

Deal with confounding/colliding

answer counterfactual questions

De-biasing via deconfounding
- ☐ Observed confounding bias
- ☐ Unobserved confounding bias

Utilize colliding structure
- ☐ Disentangle
- ☐ Model retraining

Counterfactual inference:
- ☐ (in)direct effect for debiasing
- ☐ data augmentation
- ☐ fairness
- ☐ explanation

Recommendation

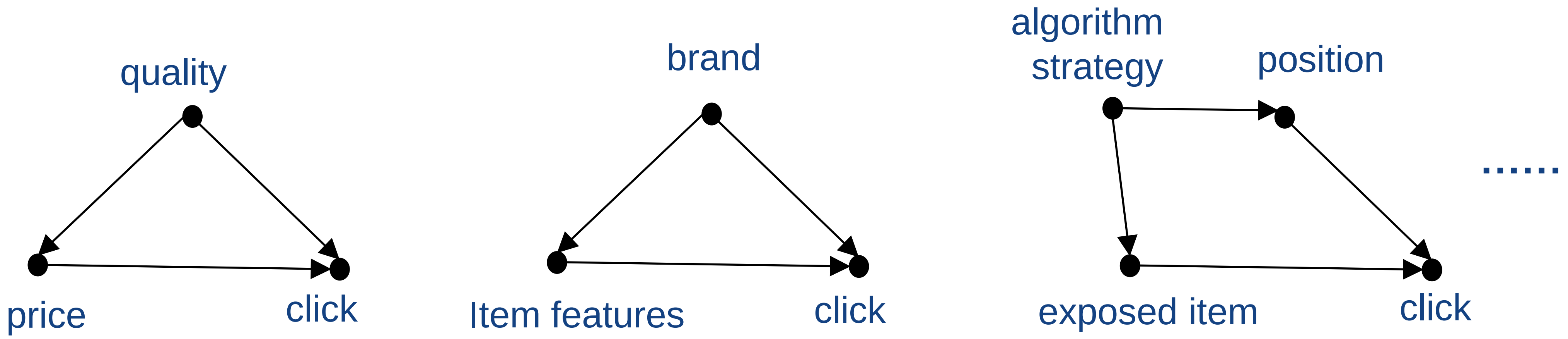
SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

Confounders in Recommendation

- Are there confounders in recommendation?

  - some examples



  - What's more, some confounder are observable/measurable, some confounder are unobservable/unmeasurable.

    e.g., company is measurable, quality is unmeasurable.

Confounders in Recommendation

- Is it necessary to deal with confounding effects?
  - The goal of recommendation: estimate user preference. But user preference is implicit.
  - We estimate it as $P(Y|U, I)$, i.e., taking the correlations between (*U*,*I*) pair and click *Y* as the preference.

- However, when there are confounders between U/I and Y(red line), the confounding effect will also bring correlations, while it cannot reflect user preference.

Thus, it is essential to deal with the confounding problem in recommendation!

But HOW?

Existing Work Regarding Observed Confounders

The backdoor adjustment is an obvious solution in this line of research.

2021 SIGIR&KDD&CIKM

- Zhang et.al. PDA
- Wang et.al. DecRS
- Yang et.al. DCM
- Gupta et.al. CauSeR

2022TKDE&KDD&CIKM&SIGIR

- Wang et.al. CaDSI
- Zhan et al. D2Q
- He et al. CISD
- Rajanala et al. DeSCoVeR

2023TOIS

- He et.al. DCR
- Zhang et.al. DML

The above work considers different problems caused by confounders, and uses different strategies to implement the backdoor adjustment.

PDA: Confounding View of Popularity Bias

- Popularity bias
  - **Favor a few popular items** while not giving deserved attention to the majority of others
  - The popular items are recommended even more frequently than their popularity would warrant, amplifying long-tail effects.
- Previous methods ignore the underline causal mechanism and blindly remove bias to purchase an even distribution.
- But, not all popularity biases data are bad.
  - Some items have higher popularity because of better quality.
  - Some platforms have the need of introducing desired bias (promoting the items that have the potential to be popular in the future).

Zhang et al. SIGIR 2021. Causal Intervention for Leveraging Popularity Bias in Recommendation

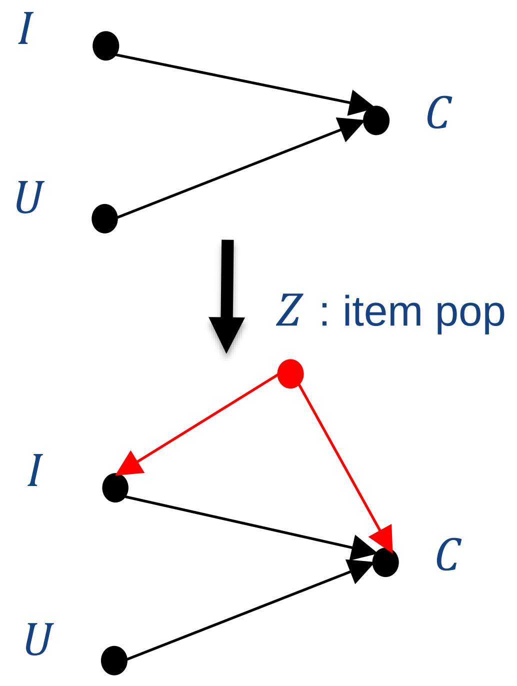
PDA: Confounding View of Popularity Bias

- What is the bad effect of popularity bias?

  - Traditional causal assumption
    - $(U, I) \rightarrow C$: user-item matching affects click.
    - Item popularity also has influence on the recommendation process, but is not considered.

  - Cofounding view
    - $Z \rightarrow I$: Popularity affects item exposure.
    - $Z \rightarrow C$: Popularity affects click probability.
    - Z is a confounder, bringing spurious (bad effect) correlation between I and C.
    - Take the causation $P(C|do(U,I))$, instead of the correlation $P(C|U,I)$, as user preference.

U: user; **I: exposed item;**
C: interaction label



Causation (backdoor adjustment):
$$\boldsymbol{P(C|do(U,I)) = \sum_Z P(C|U,I,Z)P(Z)}$$

Correlation:

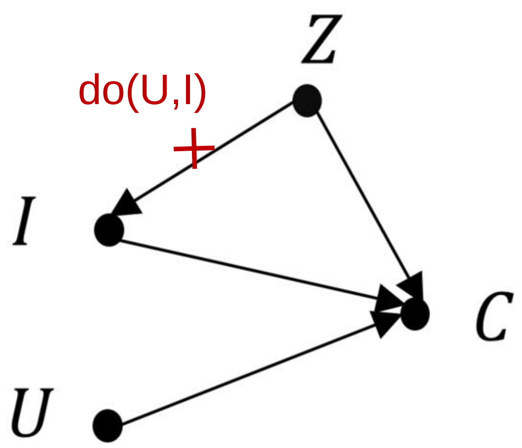
Bad effect

$$P(C|U,I) = \sum_Z P(C|U,I,Z)P(Z|I)$$
$$\propto \sum_Z P(C|U,I,Z)P(I|Z)P(Z)$$

Zhang et al. SIGIR 2021. Causal Intervention for Leveraging Popularity Bias in Recommendation

PDA: Confounding View of Popularity Bias

- **Training & Inference**: Popularity De-confounding (PD, remove bad effect)



- To estimate $P(C|do(U,I)) = \sum_z P(C|U,I,z)P(z)$
  - **Step 1**. Estimate P(C|U,I,Z)
    - $P_\Theta(c=1|u,i,m_i^t) = f_\Theta(u,i) \times m_i^t$
    - m_i^t the popularity of item i in timestamp t
    - Learn with traditional loss
  - **Step 2**. Compute $P(\mathrm{C}|\mathrm{do}(\mathrm{U},\mathrm{I}))$
    - $\sum_Z P(C|U,I,Z)P(Z) \propto f_\Theta(u,i)$
    - Derivation sees the paper

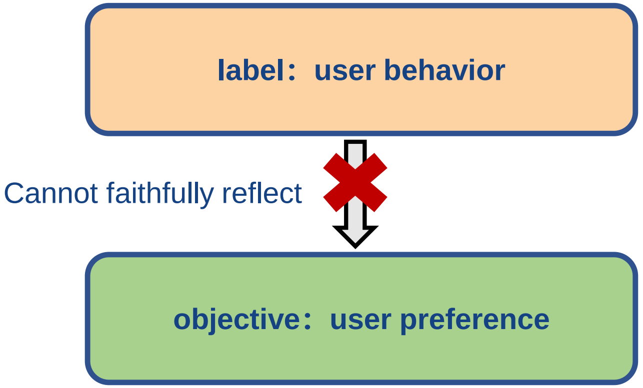
- **Another Inference**: Popularity Adjusting (inject desired popularity bias)
  - Inject the desired pop bias $\tilde{Z}$ by causal intervention

$$P(C|do(U,I),do(Z=\tilde{z})) \Longrightarrow f_\Theta(u,i) \times \tilde{m}_i$$

Zhang et al. SIGIR 2021. Causal Intervention for Leveraging Popularity Bias in Recommendation

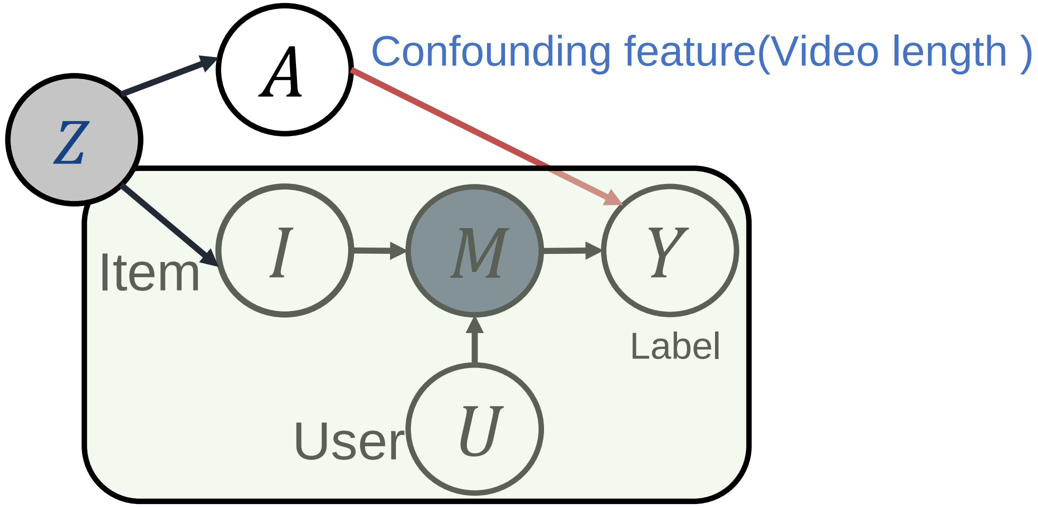
DCR: Deconfounding for Solving Unreliable Label Issue

- **Unreliable label issue:**
  - No ground-truth label for the prediction objective – user preference
  - Only have indirect label： user behaviors



progress: 60%   progress: 100%

- **Causal Modeling:**
  - Traditional assumption: U-I matching affect label
  - **Some item feature directly affect the label**

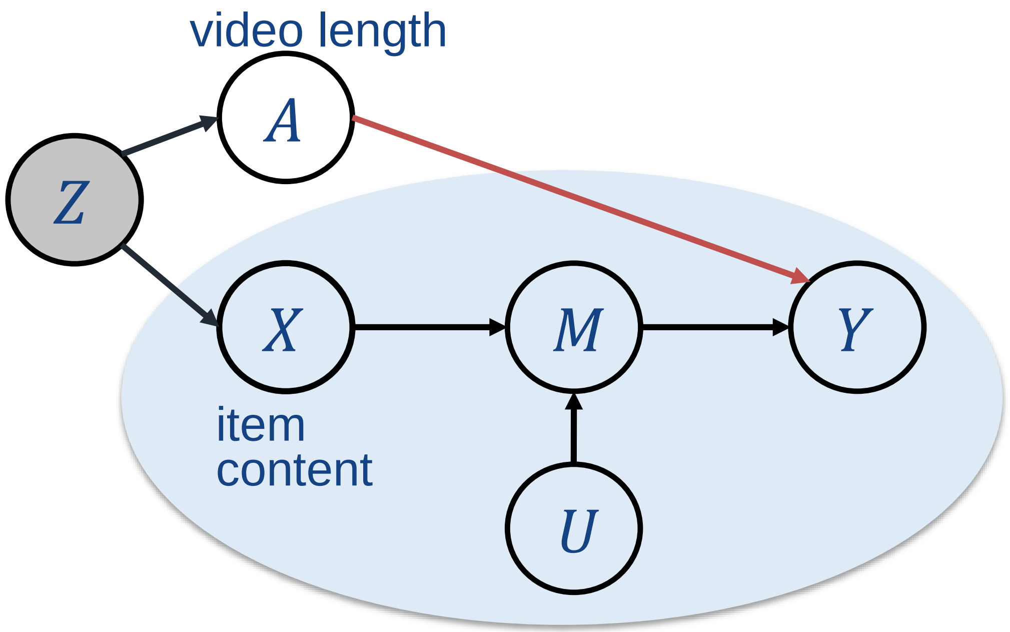


U-I matching (M) partially determines Y

He et al. Addressing Confounding Feature Issue for Causal Recommendation. TOIS 2023.

DCR: Deconfounding for Solving Unreliable Label Issue

Causal analyses



- direct path A→ Y: make P(Y|X, A) biased towards short videos
- Backdoor path $X \leftarrow Z \rightarrow A \rightarrow Y$: make P(Y|X) learn spurious correlation

Should beyond correlation-level

Causal effect as interest

true user preference: the **causal effects** path through M to Y

$$P(Y|U, do(X)) = \sum_{a \in \mathcal{A}} P(Y|U, X, A = a)P(A = a),$$

He et al. Addressing Confounding Feature Issue for Causal Recommendation. TOIS 2023.

DCR: Deconfounding for Solving Unreliable Label Issue

How to estimate the causal effect?

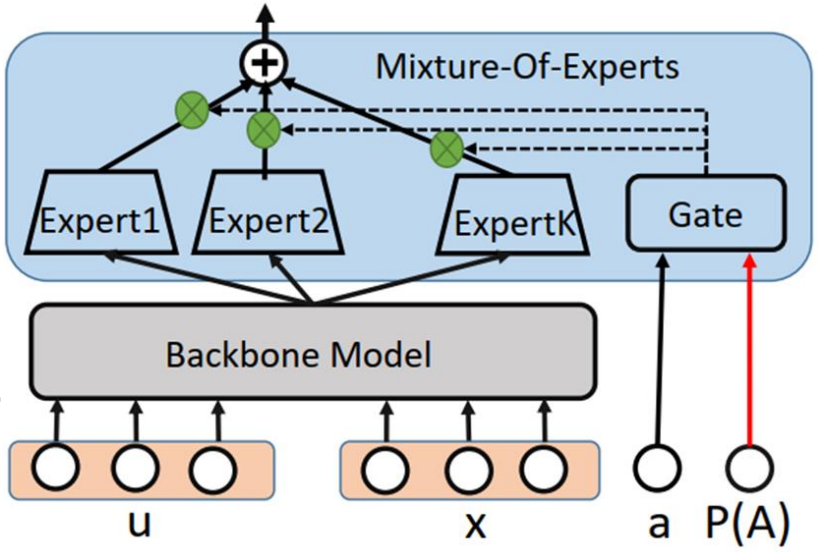
$$P(Y|U, do(X)) = \sum_{a \in \mathcal{A}} P(Y|U, X, A = a)P(A = a),$$

- **DCR: model-based estimation**

$\boxed{k^{th}\ expert:\ P(y = 1|u, x, A = a_K)}$

◆ **Training** --- fitting $P(Y|U, X, A)$

◆ **Inference** --- backdoor adjustment

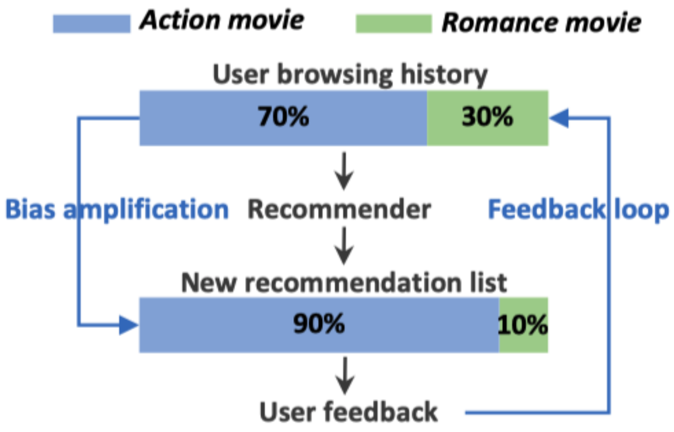


- **DCR involves changing the model architecture, DML [2] proposes to achieve the adjustment directly at the label level/**

【1】He et al. Addressing Confounding Feature Issue for Causal Recommendation. TOIS 2023.
【2】Zhang et al. Leveraging Watch-time Feedback for Short-Video Recommendations: A Causal Labeling Framework. ArXiv 2023.

DecRS: Alleviating Bias Amplification

- **Bias amplification:**
  - What is it?

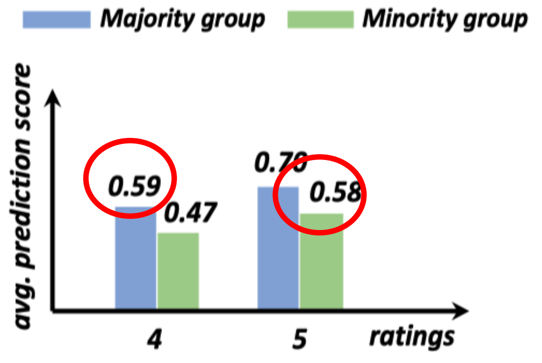


(a) An example of bias amplification.

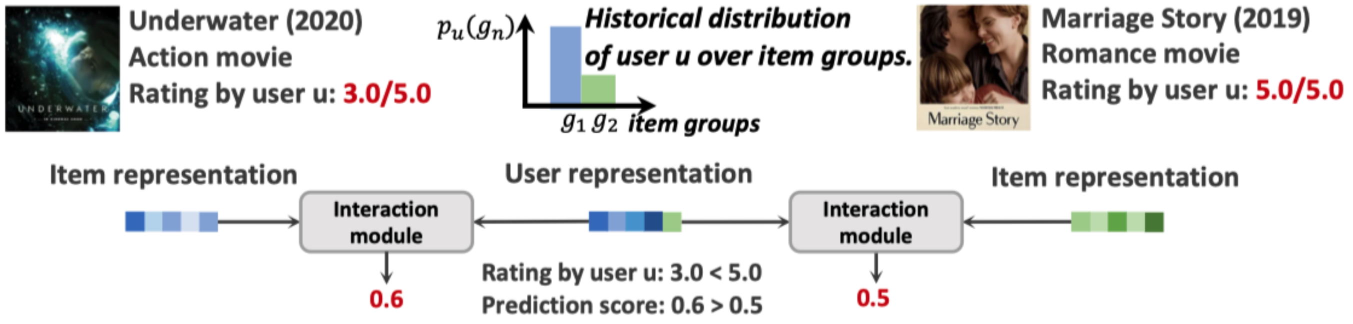
Over-recommend items in the majority group

- **Why?**

- An item with low rating receives a higher prediction score because it belongs to the majority group.
- Intuitively, we can know that the user representation shows stronger preference to majority group.



(b) Prediction score difference between the items in the majority and minority groups over ML-1M.

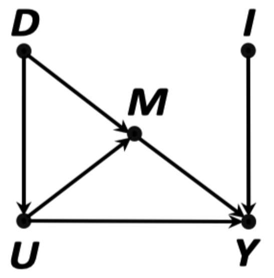


(c) An example on the cause of bias amplification.

Wang et al. SIGKDD 2021. Deconfounded recommendation for alleviating bias amplification.

DecRS: Alleviating Bias Amplification

- **Causal view of bias amplification**



(a)

U	*User representation*
I	*Item representation*
D	*User historical distribution over item groups*
M	*Group-level user representation*
Y	*Prediction score*

- D: user historical distribution over **item group.** $d_u = [p_u(g_1), \dots, p_u(g_N)]$, e.g., $d_u = [0.8, 0.2]$.
- M: describe how much **the user likes different item groups**, decided by D and U.
- $(U, M) \rightarrow Y$: an item i can have a high Y because: 1) **user's pure preference over the item** ($U \rightarrow Y$) or 2) the **user shows interest in the item group** ($U \rightarrow M \rightarrow Y$).

✓ D is a confounder between U and Y, bringing spurious correlations: given the item i in a group g, the more superior g is in u's history, the higher the prediction score Y becomes.

- **Backdoor adjustment**

$$
\begin{aligned}
& P(Y|U=\boldsymbol{u}, I=\boldsymbol{i}) \\
&= \frac{\sum_{d\in\mathcal{D}}\sum_{m\in\mathcal{M}} P(\boldsymbol{d})P(\boldsymbol{u}|\boldsymbol{d})P(\boldsymbol{m}|\boldsymbol{d},\boldsymbol{u})P(\boldsymbol{i})P(Y|\boldsymbol{u},\boldsymbol{i},\boldsymbol{m})}{P(\boldsymbol{u})P(\boldsymbol{i})} && \text{(1a)} \\
&= \sum_{\boldsymbol{d}\in\mathcal{D}}\sum_{\boldsymbol{m}\in\mathcal{M}} P(\boldsymbol{d}|\boldsymbol{u})P(\boldsymbol{m}|\boldsymbol{d},\boldsymbol{u})P(Y|\boldsymbol{u},\boldsymbol{i},\boldsymbol{m}) && \text{(1b)} \\
&= \sum_{\boldsymbol{d}\in\mathcal{D}} P(\boldsymbol{d}|\boldsymbol{u})P(Y|\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d},\boldsymbol{u})) && \text{(1c)} \\
&= P(\boldsymbol{d}_u|\boldsymbol{u})P(Y|\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d}_u,\boldsymbol{u})), && \text{(1d)}
\end{aligned}
$$

$$
\begin{aligned}
& P(Y|do(U=\boldsymbol{u}), I=\boldsymbol{i}) \\
&= \sum_{\boldsymbol{d}\in\mathcal{D}} P(\boldsymbol{d}|do(U=\boldsymbol{u}))P(Y|do(U=\boldsymbol{u}),\boldsymbol{i},M(\boldsymbol{d},do(U=\boldsymbol{u}))) && \text{(2a)} \\
&= \sum_{\boldsymbol{d}\in\mathcal{D}} P(\boldsymbol{d})P(Y|do(U=\boldsymbol{u}),\boldsymbol{i},M(\boldsymbol{d},do(U=\boldsymbol{u}))) && \text{(2b)} \\
&= \sum_{\boldsymbol{d}\in\mathcal{D}} \boxed{P(\boldsymbol{d})}P(Y|\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d},\boldsymbol{u})), && \text{(2c)}
\end{aligned}
$$

Wang et al. SIGKDD 2021. Deconfounded recommendation for alleviating bias amplification.

DecRS: Alleviating Bias Amplification

- Deconfounded Recommender System (DecRS)
  - To implement:

$$P(Y|do(U=\boldsymbol{u}),I=\boldsymbol{i})=\sum_{\boldsymbol{d}\in\mathcal{D}}P(\boldsymbol{d})P(Y|\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d},\boldsymbol{u})) \quad (3)$$

Challenge: the sample space of D is infinite.

- Backdoor adjustment approximation:

(1) Sampling distributions to represent $\mathcal{D}$;

Use function $f(\cdot)$ (FM) to calculate $\mathrm{P}(Y|u,i,M(d,u))$.

$$\begin{aligned}P(Y|do(U=\boldsymbol{u}),I=\boldsymbol{i}) &\approx \sum_{\boldsymbol{d}\in\tilde{\mathcal{D}}}P(\boldsymbol{d})P(Y|\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d},\boldsymbol{u}))\\ &=\sum_{\boldsymbol{d}\in\tilde{\mathcal{D}}}P(d)f(\boldsymbol{u},\boldsymbol{i},M(\boldsymbol{d},\boldsymbol{u})) \quad (4)\end{aligned}$$

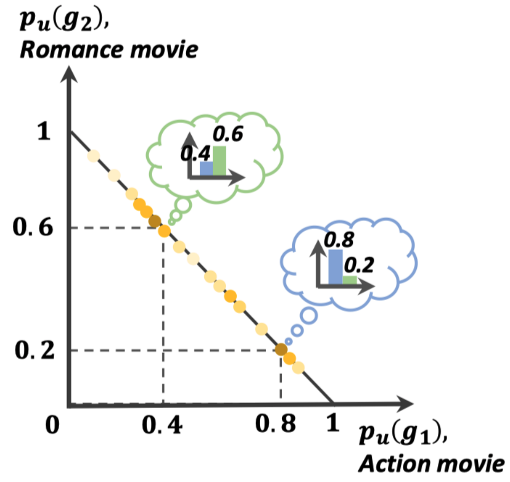
(2) Approximation of $E_d[f(\cdot)]$.

- Expectation of function $f(\cdot)$ of $\boldsymbol{d}$ in Eq. 4 is hard to compute because we need to calculate the results of $f(\cdot)$ for each $\boldsymbol{d}$.
- **Jensen's inequality**: take the sum into the function $f(\cdot)$.

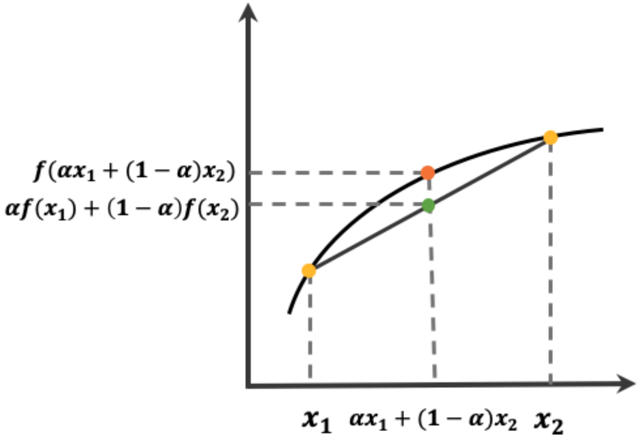
$$P(Y|do(U=\boldsymbol{u}),I=\boldsymbol{i})\approx \boxed{f(\boldsymbol{u},\boldsymbol{i},M(\sum_{\boldsymbol{d}\in\tilde{\mathcal{D}}}P(\boldsymbol{d})\boldsymbol{d},\boldsymbol{u})).} \quad (5)$$

learn it from data

Different to PDA, this term directly represents the target casual effect.



Infinite Sample Space



Approximation

Wang et al. SIGKDD 2021. Deconfounded recommendation for alleviating bias amplification.

Existing Works for Unobserved Confounders

- The methods based on backdoor adjustment need the confounders could be observable and controllable.
- However, unobserved/unmeasurable/uncontrollable confounders exist in recommendation. How to deal with them?

  - There are two lines of work:

2023 TKDE & TORS — Zhu et.al. CausalD, Xu et al. DCCF

2022 ArXiv — Zhu et.al. HCR

Front-door adjustment

2023 KDD — Zhang et.al. iDCF

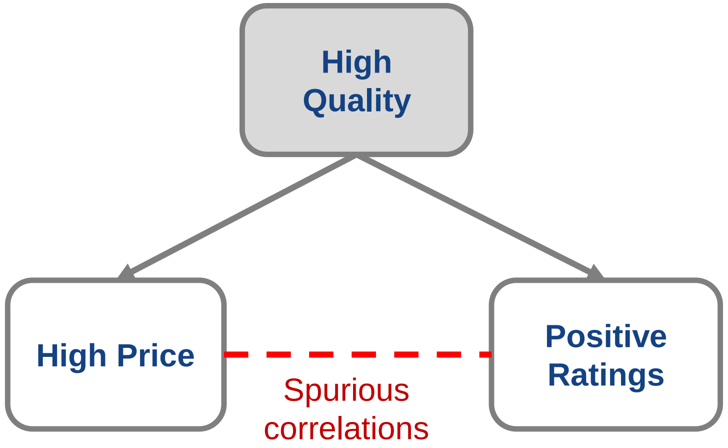
2022 ArXiv — Zhu et.al. Deep-Deconf

2020 NeurIPS, 2020 RecSys — Wang et.al. DCF, Zhou et.al. VSR

Learning substitutes

HCR: Front-door Adjustment-based Solution

- Source of confounding bias is the **confounder** that **affects item attributes and user feedback** simultaneously.
- Some confounders are hard to measure.
  - Technical difficulties, privacy restrictions, *etc*.
  - E.g., product quality.



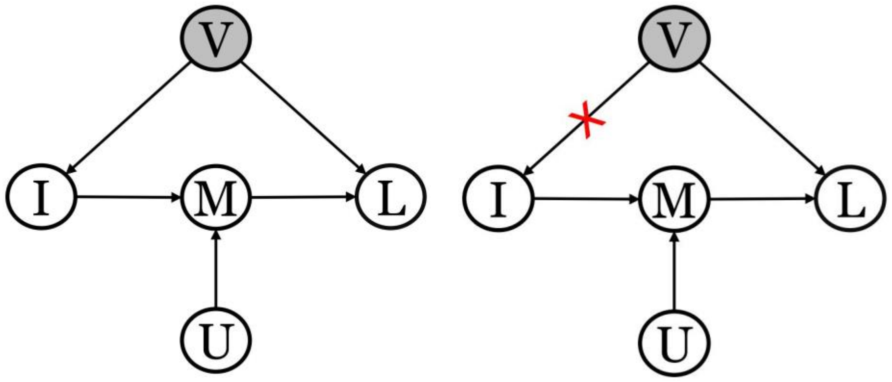
- Removing hidden confounders is hard:
  - Inverse Propensity Weighting
    - Based on strict assumption of no hidden confounder.
  - Backdoor Adjustment
    - Require the confounder’s distribution.

Xinyuan Zhu et.al. “*Mitigating Hidden Confounding Effects for Causal Recommendation*” in 2022.

HCR: Front-door Adjustment-based Solution

- Abstract user feedback generation process into causal graph.
  - V: hidden confounder; L: like feedback; I: item; U: user.
  - M: a set of variables that act as mediators between $\{U, I\}$ and L, e.g., user-item feature matching, and click.
- Key:
  - Block the backdoor path $I \leftarrow V \rightarrow L$
  - Estimate the causal effect of I on L, *i.e.*,

    $P(L|U, \mathrm{do}(I))$.



- Hidden Confounder Removal (HCR) framework.
  - Front-door adjustment
    - decompose causal effect of I on L into: 1) the effects of I on M and 2) the effect of M on L.

$$P\,(L|U, \mathrm{do}\,(I)) = \sum_M P\,(M|U, do(I))P(L|U, do(M))$$
$$= \sum_M P\,(M|U, I) \sum_{I'} P\,(I')P\,(L|M, U, I')$$

Xinyuan Zhu et.al. “*Mitigating Hidden Confounding Effects for Causal Recommendation*” in 2022.

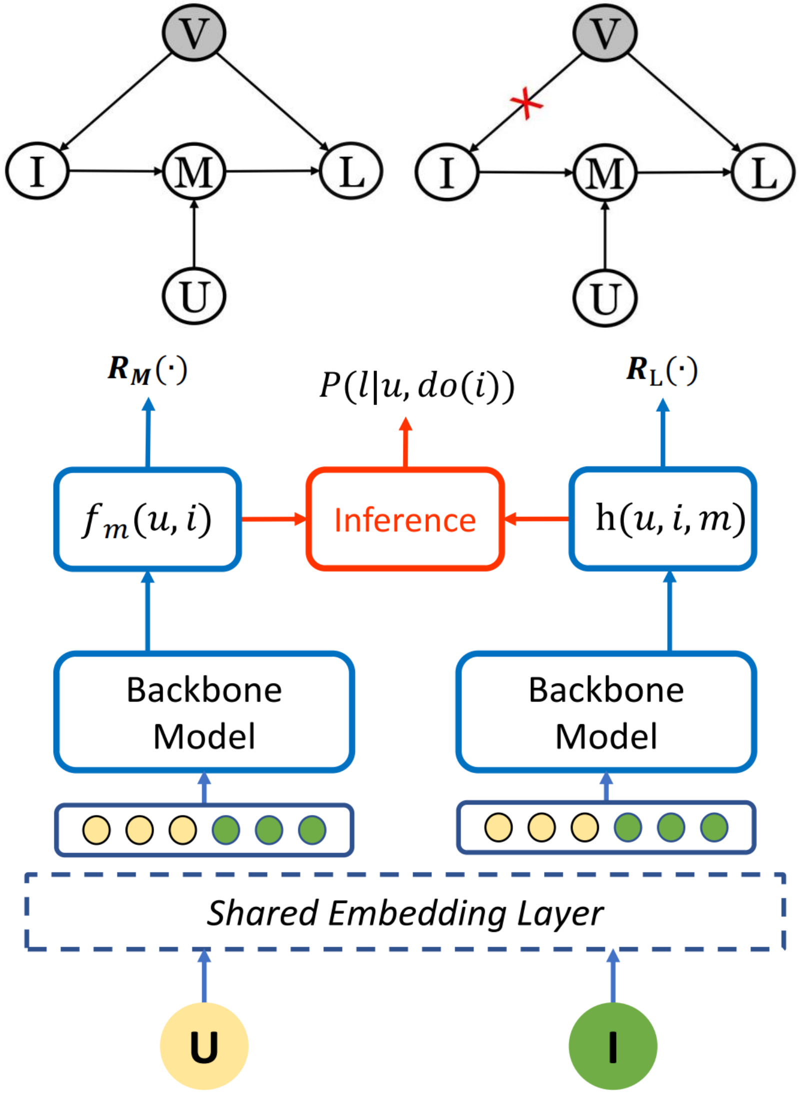
HCR: Front-door Adjustment-based Solution

- Hidden Confounder Removal (HCR) framework
  - $P(L|do(I),U) = \sum_M P\ (M|U,I) \sum_{I'} P\ (I')P\ (L|U,I',M)$
  - Multi-task learning
    - Learns $P(M|U,I) \coloneqq f_m(U,I)$
    - Learn

$$\begin{aligned}\mathrm{P}(L|M,U,I) &\coloneqq h(U,I,M)\\ &= h^1(U,M)h^2(U,I')\end{aligned}$$

  - Inference
    - Infer $P(M|U,I)$ and $P(L|U,I,M)$
    - Get rid of the sum over I and obtain

$$\begin{aligned}&P(L|U,do(I))\\ &= \sum_M f_m(U,I) \sum_{I'} P(I')h^1(U,M)h^2(U,I')\\ &= \sum_M f_m(U,I)h^1(U,M) \sum_{I'} P(I')h^2(U,I')\\ &= S_u \sum_M f_m(U,I)h^1(U,M)\end{aligned}$$



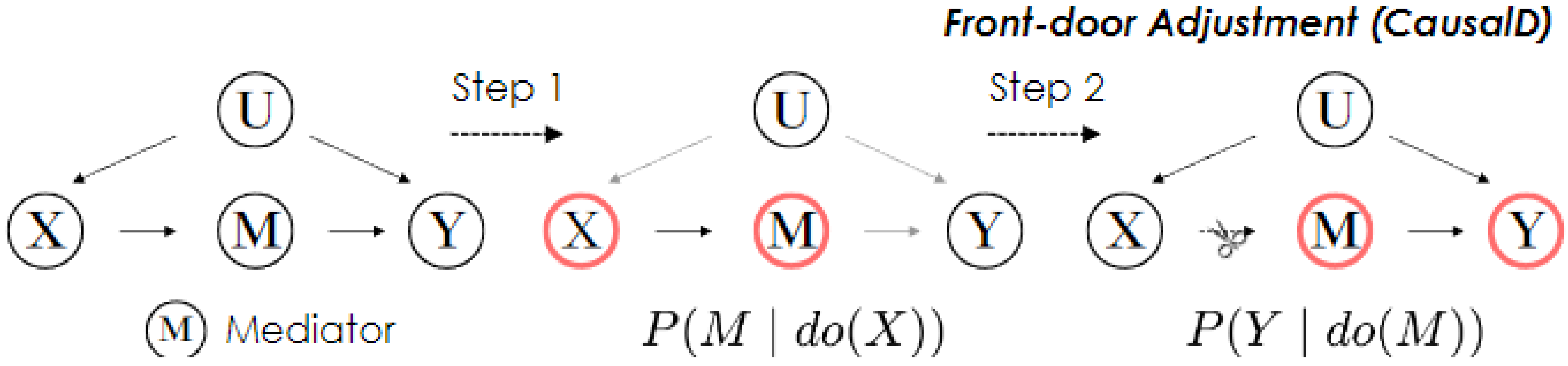
Xinyuan Zhu et.al. "*Mitigating Hidden Confounding Effects for Causal Recommendation*" in 2022.

CausalD: Front-door Adjustment-based Solution

- Consider Hidden Confounder in Sequential Recommendation

Sequential recommendation: predict user next behavior using historical behaviors

X: historical interaction Y: Next behavior M: Representations
U: unobserved confounder, such as social relationships



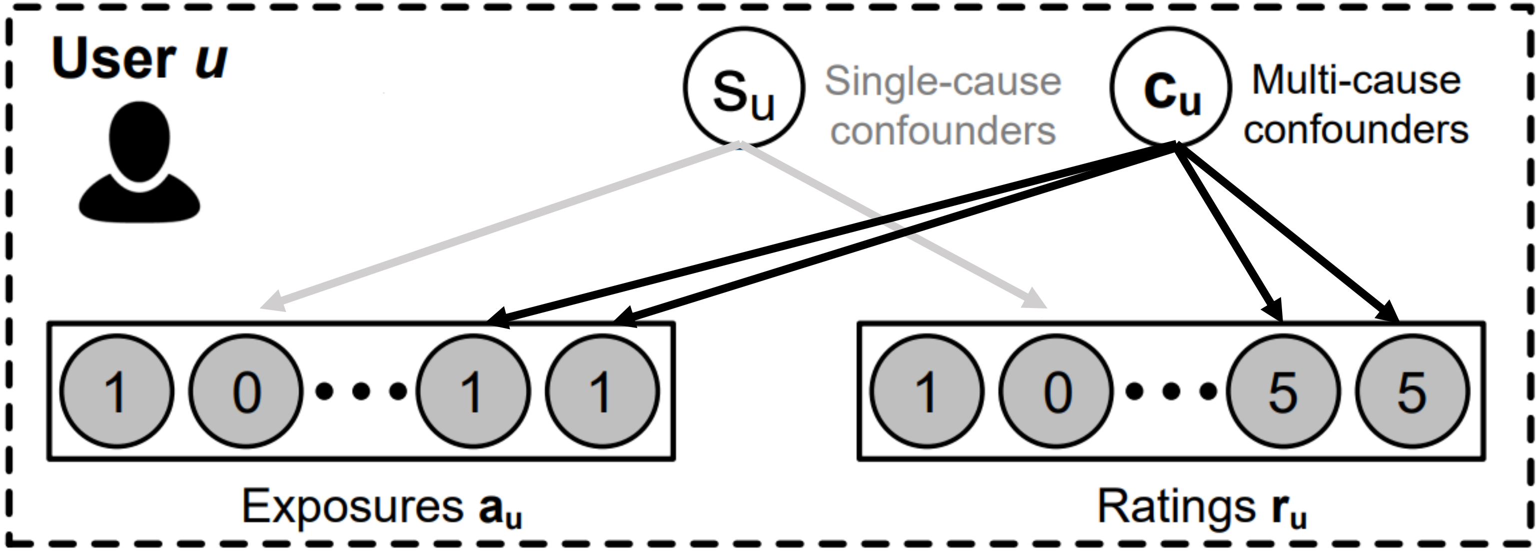
$$P\,(Y|\mathrm{do}\,(X)) = \sum_m P\,(m|do(X))P(Y|do(m))$$
$$= \sum_m P\,(m|X) \sum_{x'} P\,(X = x')P\,(Y|m, x')$$

- Estimation method: similar to HCR but in a distillation manner

Zhang et.al. "Causal Distillation for Alleviating Performance Heterogeneity in recommender System" in TKDE 2023.

Learning Substitutes-based Solution

- Multiple causes assumption for recommendation:
  - multiple causes: each user's binary exposure to an item a_{ui} is a cause(treatment), thus there are multiple causes.
  - There are multiple-cause confounders (confounders that affect ratings and many causes).
  - Single-cause confounders (confounders that affect ratings and only one cause) are negligible.



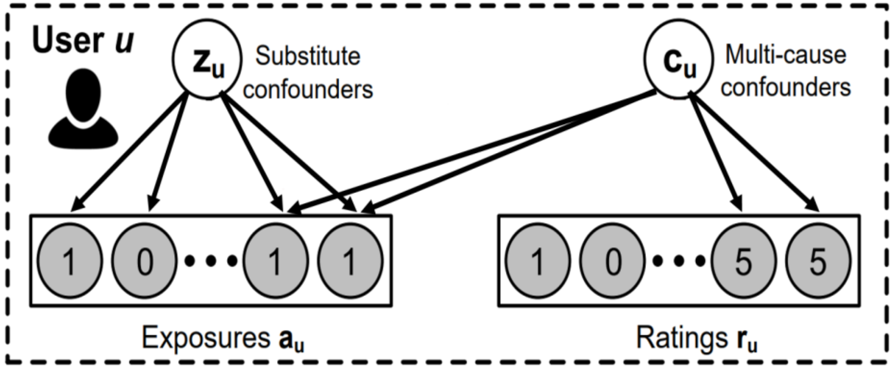
Wang et al. RecSys 2020. Causal inference for recommender system.
Zhu et.al. Arxiv 2022. Deep causal reasoning for recommendations.

Learning Substitutes-based Solution

- Learning substitutes to deconfounding:

Key: if Z_u renders the $a_{u,i}$'s conditionally independent then there cannot be another multi-cause confounder

Contradiction: assume $p(a_{u1}, \ldots, a_{um} | z_u) = \prod_i p(a_{ui} | z_u)$, if there is a multi-cause confounder, the conditional independence cannot hold.



- Step 1: learning substitutes

Finding a Z_u, such that:

$$p(a_{u1}, \ldots, a_{um} | z_u) = \prod_i p(a_{ui} | z_u)$$

Example:

find a generative model:

$$P_\Theta(A_u | Z_u) = \prod_{i=1}^{m} Bern(a_{ui} | \theta(z_u)_i)$$

then:

find $q_\Phi(Z_u | A_u)$ with variation-inference

- Step 2: deconfounded recommender

Control the substitutes to fit recommender model

Example:

$$y_{ui}(a) = \theta_u^\top \beta^i \cdot a + \gamma_u \cdot z_{ui} + \epsilon_{ui}$$

where θ_u and β_i refer user preference and item attributes, respectively.

Wang et al. RecSys 2020. Causal inference for recommender system.
Wang et.al. *J Am Stat Assoc 2019.* The blessings of multiple causes.
Zhu et.al. Arxiv 2022. Deep causal reasoning for recommendations.

Papers on Deconfounding Recommendation

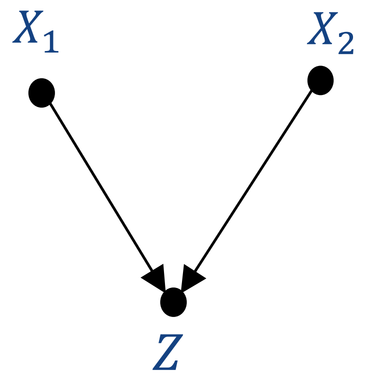
- Zhang, Yang, et al. "Causal intervention for leveraging popularity bias in recommendation." In *SIGIR* 2021. (Zhang et.al. PDA)
- Wang, Wenjie, et al. "Deconfounded recommendation for alleviating bias amplification." In *SIGKDD* 2021. (wang et.al. DecSR)
- Wang, Xiangmeng, et al. "Causal Disentanglement for Semantics-Aware Intent Learning in Recommendation." In *TKDE* 2022. (Wang et.al. CaDSI)
- Gupta, Priyanka, et al. "CauSeR: Causal Session-based Recommendations for Handling Popularity Bias." In *CIKM* 2021. (Gupta et.al., CauSeR)
- Rajanala, Sailaja, et al. "Descover: Debiased semantic context prior for venue recommendation." In *SIGIR* 2022 (Rajanala et al. DeSCoVeR)
- Yang, Xun, et al. "Deconfounded video moment retrieval with causal intervention." In *SIGIR 2021*. (Yang et.al. DCM)
- Zhan, Ruohan, et al. "Deconfounding Duration Bias in Watch-time Prediction for Video Recommendation." SIGKDD 2022. (Zhan et al. D2Q)
- He, Ming, et al. "Causal intervention for sentiment de-biasing in recommendation." *In CIKM 2022*. (He et al. CISD)
- He, Xiangnan, et al. "Addressing confounding feature issue for causal recommendation." *ACM TOIS* 2023. (He et al. DCR)
- Wang, Yixin, et al. "Causal inference for recommender systems." Fourteenth ACM Conference on Recommender Systems. 2020. (Wang et.al. DCF)
- Zhang, Yang, et al. "Leveraging Watch-time Feedback for Short-Video Recommendations: A Causal Labeling Framework." arXiv 2023. (Zhang et al. DML)
- S. Zhang *et al.*, "Causal Distillation for Alleviating Performance Heterogeneity in Recommender Systems," TKDE 2023. (Zhang et al. CausalD)
- Qing Zhang et.al. Debiasing Recommendation by Learning Identifiable Latent Confounders. KDD 2023. (Zhang et al. iDCF)
- Zhu, Xinyuan, et al. "Mitigating hidden confounding effects for causal recommendation." *arXiv 2022. (Zhu et al. HCR)*

SCM for Recommendation

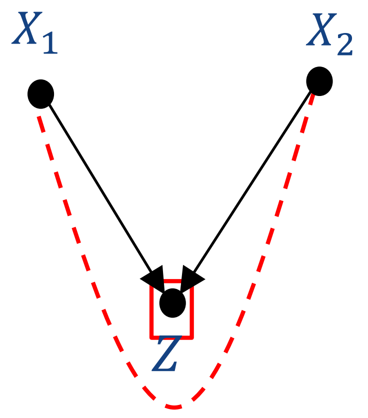
- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

Colliding Effects in Recommendation

- Are there colliders in recommendation?
  - There are variables affected by many factors. Such as, the happening of clicking is affected by user preference and the exposure position.
  - Existing work also tries to construct colliders manually.



- To utilize or eliminate colliding effects?
  - Assume that we have known X_2, try to estimate X_1.
  - Condition on Z, X_1 and X_2 could be correlated.
  - That means condition on Z, X_2 would provide us more information to estimate X_1.



In recommendation, we usually face with this case (know X_2 and Z to predict X_1). Thus existing work based on SCM tries to utilize colliding effects to better learn some targets.

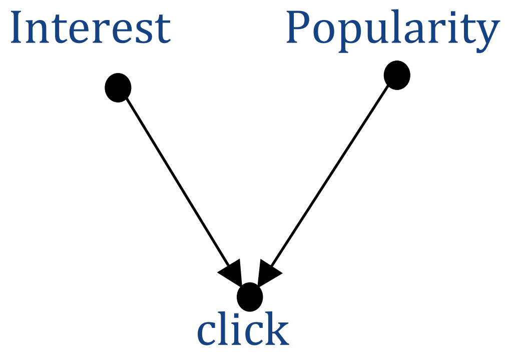
DICE: Colliding Effects for Disentangling True Interest

- What are causes of a user-item interaction (click)?

  Two main causes:
  - Interest
  - Conformity
  User tend to follow the mainstream

  - Disentangle Interest and Conformity to identify true interest.
  - But it is hard because of lacking ground-truth. (An interaction can come from either factor or both factors)

- Colliding effect can come to help:

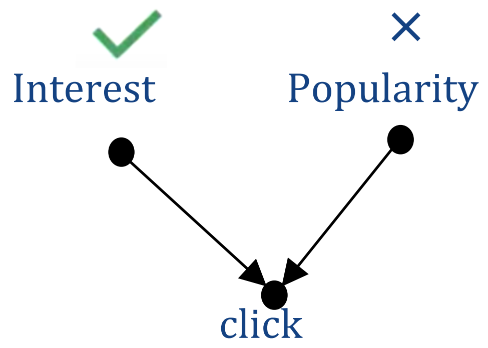


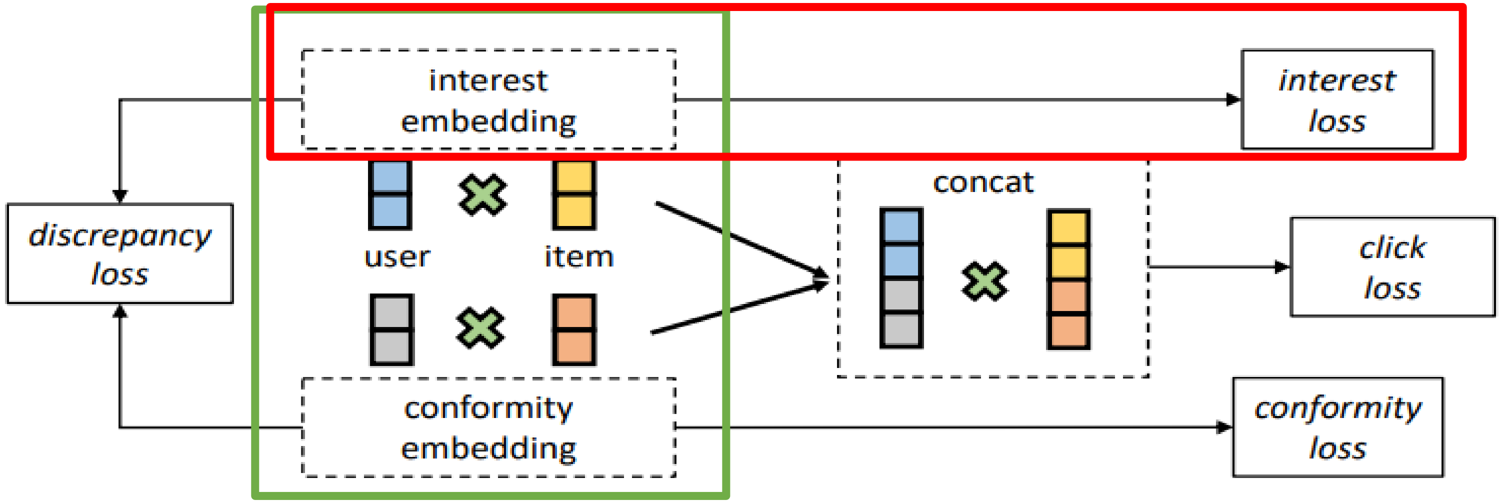
  - Interest and Popularity (conformity) are independent
  - But, they are correlated given clicks:
  A click on less popular item → High Interest

Zheng et al. WWW 2021. Disentangling User Interest and Conformity for Recommendation with Causal Embedding

DICE: Colliding Effects for Disentangling True Interest

- DICE: Partial pairwise data identifies true interest:
  - O_1: *{<u, pos_item, neg_item>}*, wherein *pos_item* is less popular than *neg_item}*
  - Pairwise cause-specific data (interst-driven): we can ascertain that the interaction is more likely due to user interest





- **Key1: split user/item representation into two embeddings**

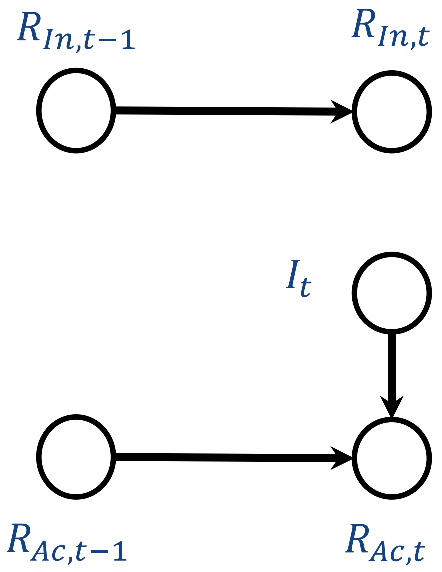
- **Key 2: learning interest embedding on interest-driven pairwise data (O_1).**

- The core idea of leveraging colliding effects has also been extended to Sequential Recommendation. (Sun et al. MiceRec. 2022.)

Zheng et al. Disentangling User Interest and Conformity for Recommendation with Causal Embedding. In WWW 2021.
Sun et al. Multi-interest Sequence Modeling for Recommendation with Causal Embedding. In SDM 2022. (Sun et al. MiceRec.)

Colliding Effects for Incremental Training

- Incremental training for recommender system
  - Usually, using the incremental interaction data I_t for efficient retraining.
  - Only updating the representations of **active** user/item corresponding to I_t.
  - Ignoring the representations of **inactive** user/item.



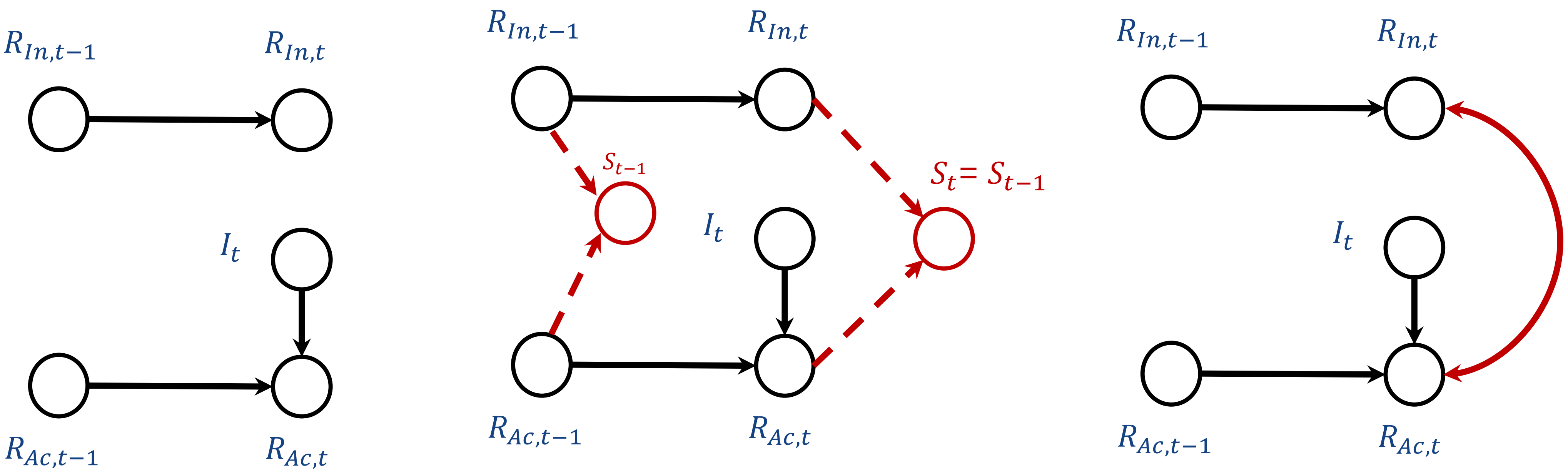
Causal graph of incremental training

- $R_{In,t-1}$: Representations of inactivate user/item at time *t-1*.
- $R_{In,t}$: Representations of inactivate user/item at time *t*.
- $R_{Ac,t-1}$: Representations of activate user/item at time *t-1*.
- $R_{Ac,t}$: Representations of activate user/item at time *t*.
- I_t: Incremental interaction data collected from time *t-1* to *t*.

Ding, Sihao, et al. "Causal incremental graph convolution for recommender system retraining." IEEE TNNLS (2022).

Colliding Effects for Incremental Training

- Causal incremental training with colliding effects



Building colliding effect

- Creating a collider S_t between $R_{In,t}$ and $R_{Ac,t}$, S_t is the similarity between representations of active and inactivate user/item.
- Restraining $S_t = S_{t-1}$ to open the causal path $I_t \rightarrow R_{Ac,t} \rightarrow R_{In,t}$ with the help of colliding effect.
- Using the incremental data I_t simultaneously update both $R_{Ac,t}$ and $R_{In,t}$.

Ding, Sihao, et al. "Causal incremental graph convolution for recommender system retraining." IEEE TNNLS (2022).

SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

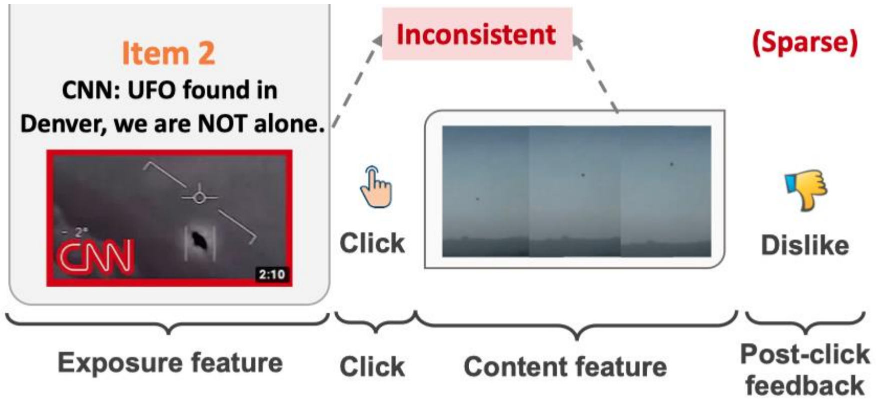
Counterfactual Recommendation

- Counterfactual inference for debiasing
  - Focus on **removing path-specific effects** for debiasing
  - First estimate the causal effect by comparing a counterfactual world with the factual world, and then mitigate path-specific effects.

- Representative Work
  - Wang, et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In SIGIR 2021.
  - Wei, et al. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system. In KDD 2021.
  - Zihao Zhao et al. Popularity Bias Is Not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. In TKDE (2022).
  - Gang Chen et al. Unbiased Knowledge Distillation for Recommendation. In WSDM 2023.

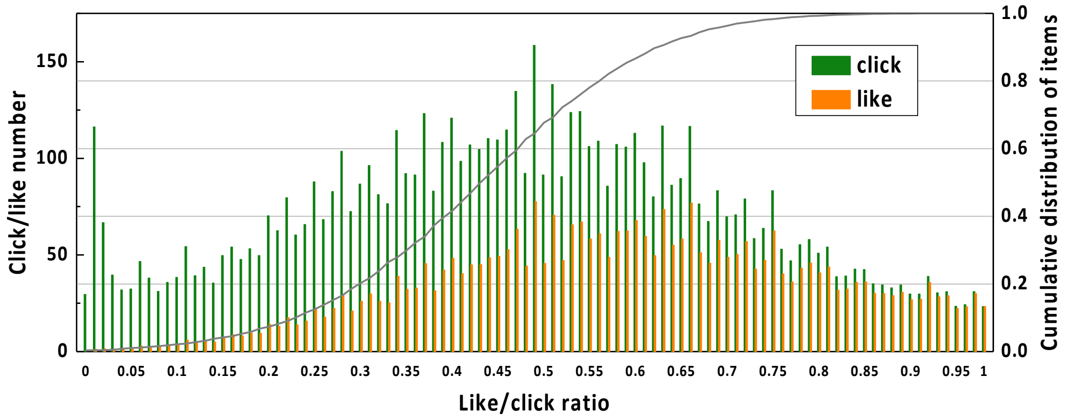
Counterfactual for Mitigating Clickbait Bias

- **Clickbait bias**
  - User interactions are biased to the items with attractive exposure features.
  - Clickbait items: exposure features (*e.g.*, title/cover image) attract users while content features (*e.g.*, video) are disappointing.
  - Recommender models learned from the biased interactions will frequently recommend these clickbait items, decreasing user experience.



(a) Browsing behaviors of users on Tiktok.

Fig. Statistics of clicks and likes on Tiktok dataset. Partly show the wide existence of clickbait issue.



Wang et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In SIGIR 2021.

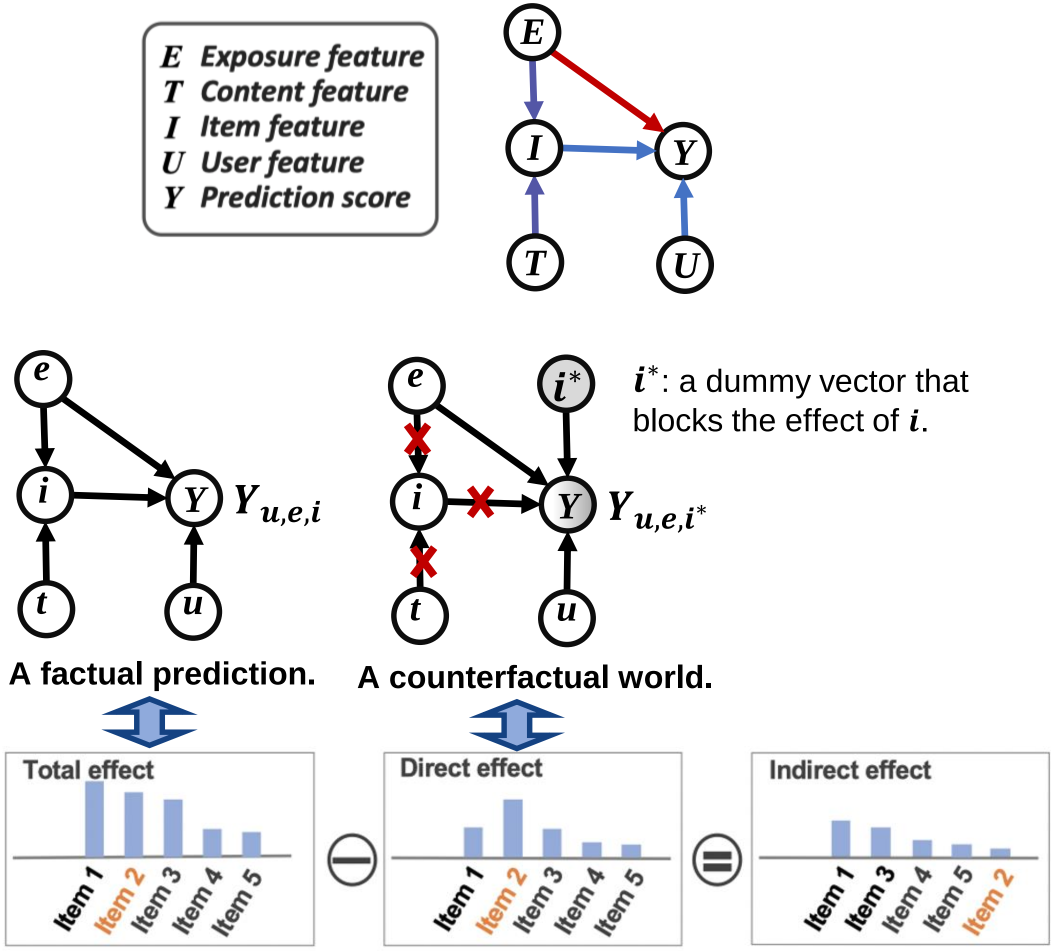
Counterfactual for Mitigating Clickbait Bias

- ## Counterfactual Inference

❖ **Causal Graph**
- A causal graph to describe the causal relationships between the features and user-item prediction scores.
- **Reason for clickbait issue:** $E \rightarrow Y$ a clickbait item has high prediction scores purely due to its attractive exposure features, *i.e.*, title/cover.

❖ **Causal learning** for training**:** learn structural functions $I(E,T)$ and $Y(U,I,E)$ from data.

❖ **Causal reasoning** for inference**:** counterfactual inference.
- Reduce the direct effect of exposure features.
- **1) Estimate** the effect in the counterfactual world, which imagines *what the prediction score would be if the item had only the exposure features.*
- **2) Reduce** the direct effect of exposure features for inference.



Wang et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In SIGIR 2021.

Counterfactual for Mitigating Clickbait Bias

- **Overall Performance**

Table 2: Top-K recommendation performance of compared methods on Tiktok and Adressa. %Improve. denotes the relative performance improvement of CR over NT. The best results are highlighted in bold. Stars and underlines denote the best results of the baselines with and without using additional post-click feedback during training, respectively.

Dataset	Tiktok						Adressa					
Metric	P@10	R@10	N@10	P@20	R@20	N@20	P@10	R@10	N@10	P@20	R@20	N@20
NT [50]	<u>0.0256</u>	<u>0.0357</u>	0.0333	<u>0.0231</u>	<u>0.0635</u>	0.0430	<u>0.0501</u>	<u>0.0975</u>	<u>0.0817</u>	<u>0.0415</u>	<u>0.1612</u>	<u>0.1059</u>
CFT [50]	0.0253	0.0356	<u>0.0339</u>	0.0226	0.0628	<u>0.0437</u>	0.0482	0.0942	0.0780	0.0405	0.1573	0.1021
IPW [27]	0.0230	0.0334	0.0314	0.0210	0.0582	0.0406	0.0419	0.0804	0.0663	0.0361	0.1378	0.0883
CT [50]	0.0217	0.0295	0.0294	0.0194	0.0520	0.0372	0.0493	0.0951	0.0799	0.0418*	0.1611	0.1051
NR [51]	0.0239	0.0346	0.0329	0.0216	0.0605	0.0424	0.0499	0.0970	0.0814	0.0415	0.1610	0.1058
RR	0.0264*	0.0383*	0.0367*	0.0231*	0.0635*	0.0430*	0.0521*	0.1007*	0.0831*	0.0415	0.1612*	0.1059*
CR	**0.0269**	**0.0393**	**0.0370**	**0.0242**	**0.0683**	**0.0476**	**0.0532**	**0.1045**	**0.0878**	**0.0439**	**0.1712**	**0.1133**
%Improve.	5.08%	10.08%	11.11%	4.76%	7.56%	10.70%	6.19%	7.18%	7.47%	5.78%	6.20%	6.99%

- **Observations:**
  - **RR** achieves the best performance in the baselines by using post-click feedback for reranking.
  - **Proposed CR** significantly recommends more satisfying items by mitigating clickbait bias.

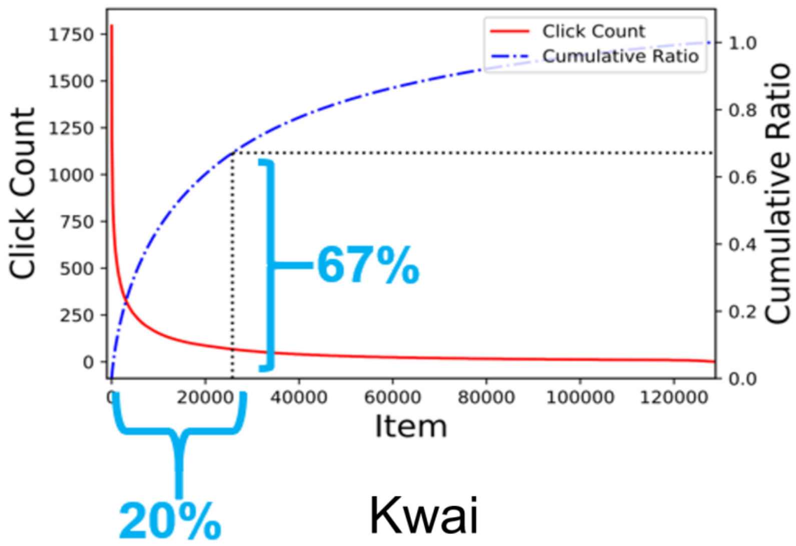
Wang et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In SIGIR 2021.

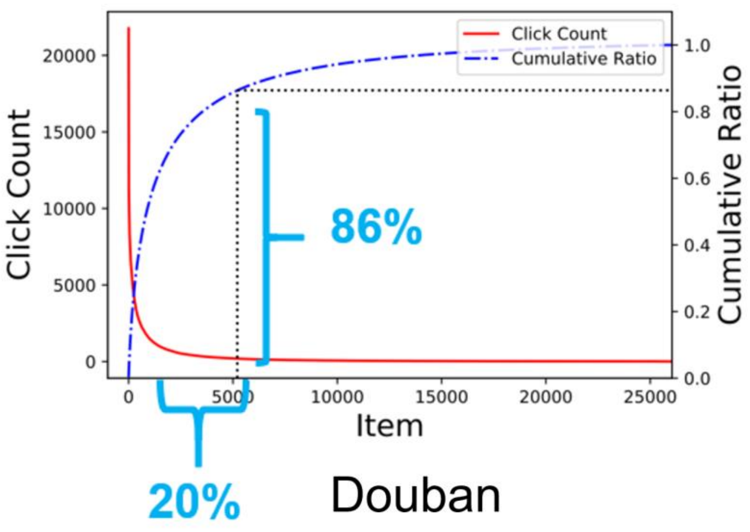
Counterfactual for Mitigating Popularity Bias

Popularity Bias in RecSys

- Popularity bias ≠ Uneven popularity distribution
  - The popular items are gradually over-recommended, amplifying long-tail effects.
  - Favor a few popular items while not giving deserved attention to the majority of others.

- From data perspective:



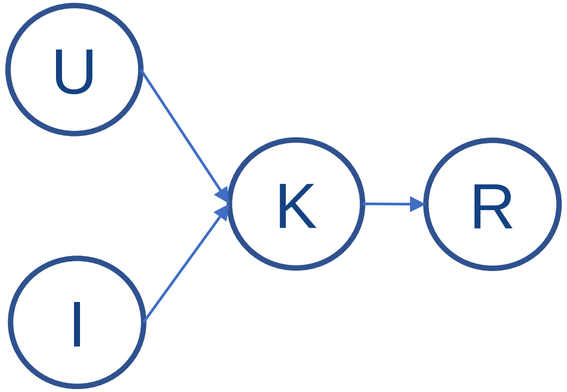


Long-tail distribution

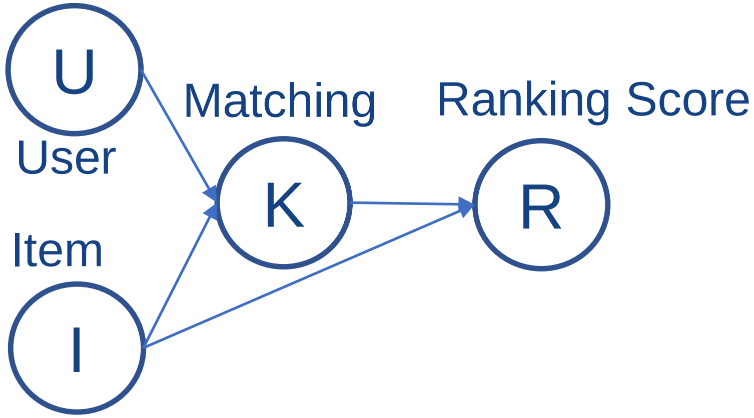
Wei et al. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In KDD 2021.

Counterfactual for Mitigating Popularity Bias

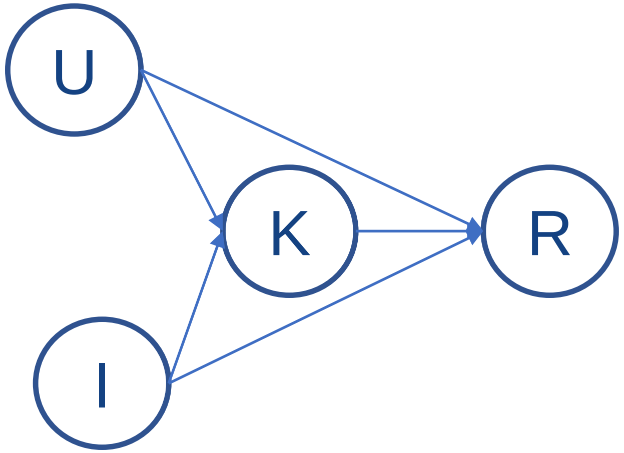
- Causal View of Popularity Bias



Common Recommender
User-Item Matching



Popularity bias modeling:
Incorporating item popularity



User-specific modeling:
Incorporating item popularity & user activity

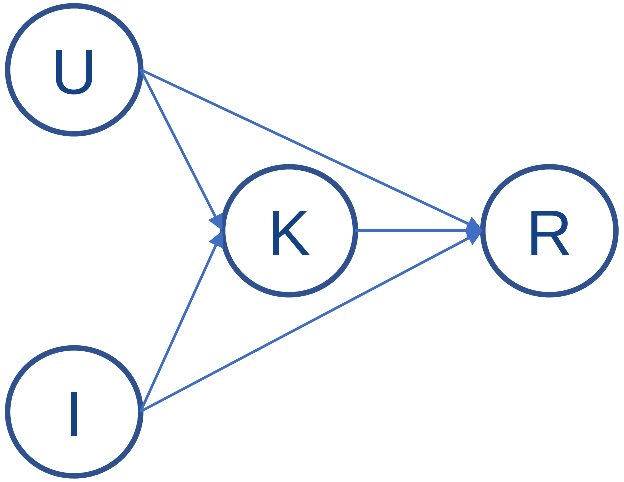
  - Edge I→R captures popularity bias.
  - Edge U→R captures the user sensitive to popularity.

- Solution:
  - Train a recommender based on the causal graph via a multi-task learning
  - Perform counterfactual inference to eliminate popularity bias (*Question to answer: what would the prediction be if there were only popularity bias?*)

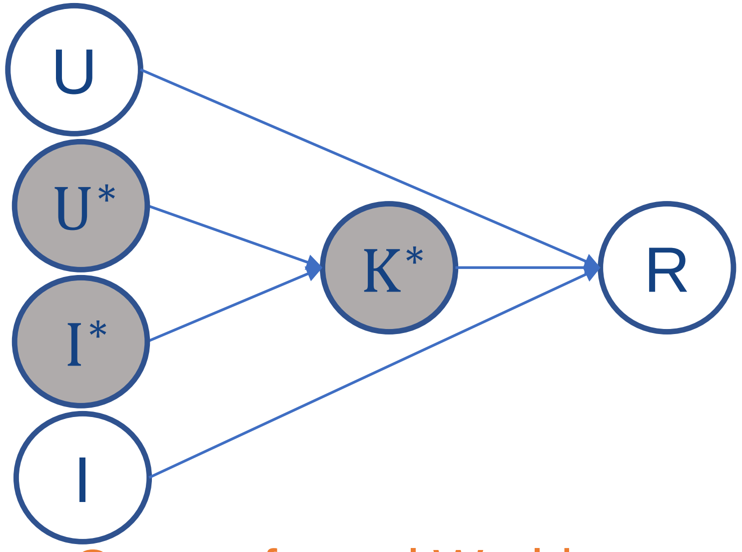
Wei et al. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In KDD 2021.

Counterfactual for Mitigating Popularity Bias

- **Counterfactual Inference to Remove Bias**
- *Question: what the prediction would be if there were no bias?*



Factual World
(original prediction)



Counterfactual World
(block matching to capture bias)

$$TIE = TE - NDE = Y(U = u, I = i, K = K_{u,i}) - Y(U = u, I = i, K = K_{u^*,i^*})$$

Factual world   Counterfactual world

Inference with TIE $= \hat{y}_k \times \sigma(\hat{y}_i) \times \sigma(\hat{y}_u)$ - $c \times \sigma(\hat{y}_i) \times \sigma(\hat{y}_u)$

Wei et al. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In KDD 2021.

Counterfactual for Mitigating Popularity Bias

- Evaluate MACR framework on two base models: MF and LightGCN.
- Testing data is intervened to be uniform.

MF as the backbone

data / Method	Adressa		Yelp2018	
	Recall	NDCG	Recall	NDCG
MF	0.0853	0.0341	0.0060	0.0094
ExpoMF	0.0896	0.0365	0.0060	0.0093
MF_causE	0.0835	0.0365	0.0051	0.0083
MF_BS	0.0900	0.0377	0.0061	0.0098
MF_reg	0.0659	0.0332	0.0050	0.0081
MF_IPS	**0.0964**	**0.0392**	**0.0062**	**0.0100**
MACR	0.1090	0.0495	0.0264	0.0192

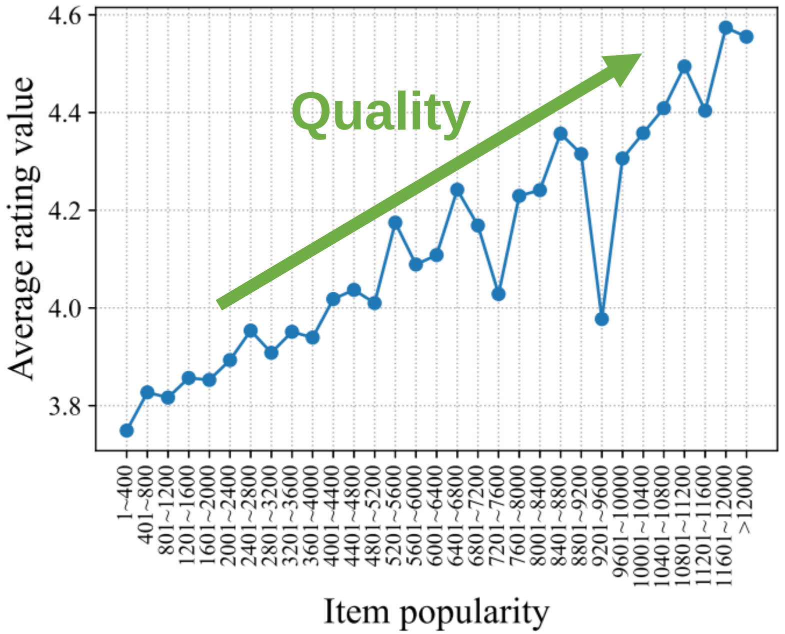
LightGCN as the backbone

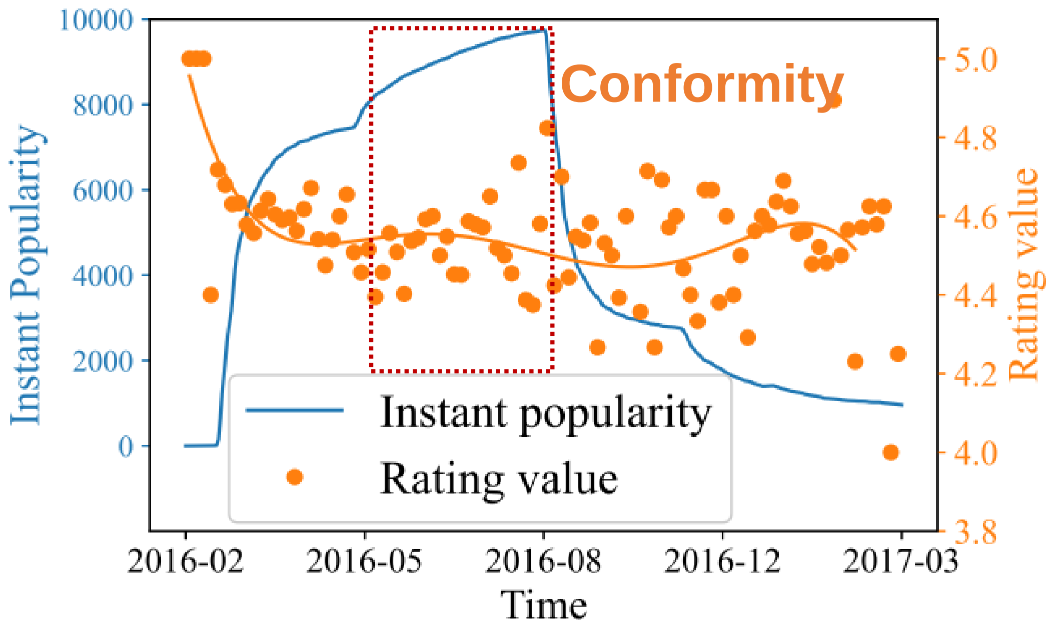
data / Method	Adressa		Yelp2018	
	Recall	NDCG	Recall	NDCG
Lgcn	0.0977	0.0395	0.0044	0.0086
Lgcn_causE	0.0823	0.0374	0.0050	0.0088
Lgcn_BS	0.1085	0.0469	**0.0048**	**0.0088**
Lgcn_reg	**0.0979**	**0.0390**	0.0042	0.0083
Lgcn_IPS	0.1070	0.0468	0.0054	0.0090
MACR	0.1273	0.0525	0.0312	0.0177

Wei et al. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In KDD 2021.

Counterfactual for Leveraging Popularity Bias

- **Conflicting Observation:**
  - The more **popular** an item is, the larger average **rating value** the item tends to have (**positive** correlation).
  - From the temporal view, for a large proportion of items, the **rating value** exhibits **negative** correlation with the item **popularity** at that time
- Quality + Conformity → Popularity, thus disentangle benign and harmful Bias





Zhao et al. Popularity Bias is not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. TKDE' 22.

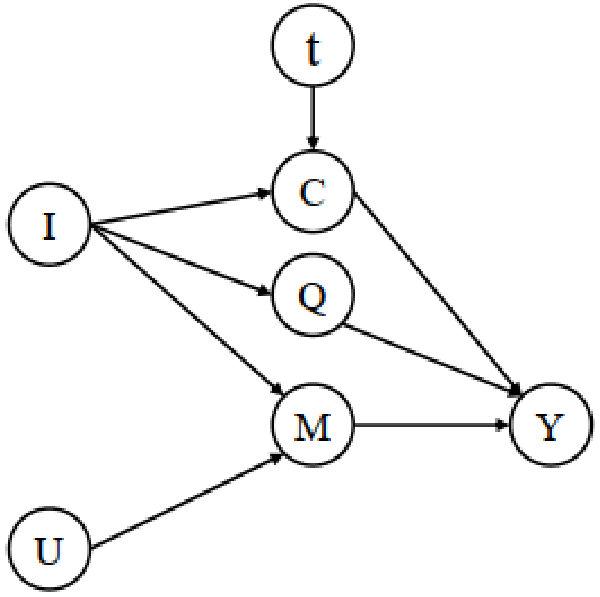
Counterfactual for Leveraging Popularity Bias

- **T**ime-aware **D**is**E**ntangled framework(**TIDE**)
  - Main challenge: Lack of explicit signal for disentanglement

- **Quality** is **static:** $\mathrm{I} \rightarrow Q \rightarrow Y$
  - Quality has **stable** influence on users' behavior

- **Conformity** is **dynamic:** $(\mathrm{I}, \mathrm{t}) \rightarrow C \rightarrow Y$
  - Conformity is **time-sensitive**

- **User interest:** $(\mathrm{U}, \mathrm{I}) \rightarrow M \rightarrow Y$
  - User and item's matching score, can be Implemented by various recommendation models, such as MF, LightGCN, etc.



(a) Causal graph of our TIDE.

U: User I: Item
t: time C: conformity
Q: Quality Y: Prediction
M: Matching score

Zhao et al. Popularity Bias is not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. TKDE' 22.

Counterfactual for Leveraging Popularity Bias

- **Training Stage**:
  - **Popularity** comes from **Quality** and **Conformity**
  - Prediction with **Popularity** and **matching score**

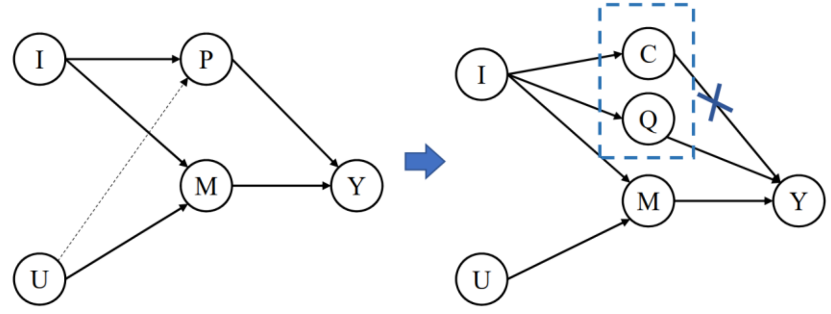
$$\hat{y}_{ui}^{t} = \mathrm{Tanh}(q_i + c_i^t) \times \mathrm{Softplus}(m_{ui})$$

- **Inference Stage**:
  - Intervention: set c as reference vector c^* (e.g., zero) during inference to **remove** the **improper effect from C to Y**.

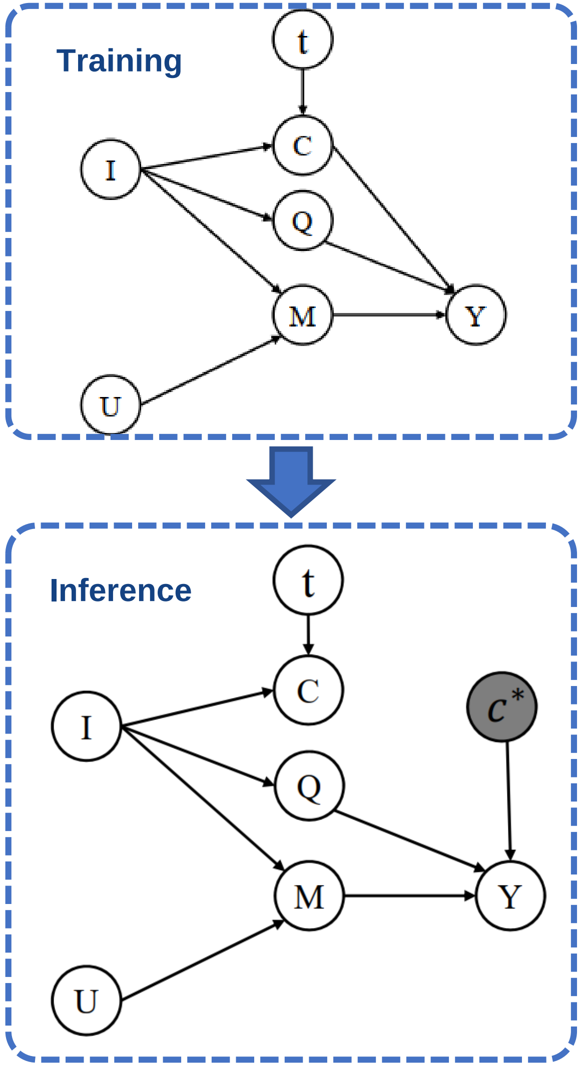
$$\hat{y}_{ui}^{*} = \tanh(q_i + c^*) \times \mathrm{Softplus}(m_{ui})$$

- Comparison with PD
  - TIDE further conduct disentanglement of popularity bias

PD



TIDE



Zhao et al. Popularity Bias is not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. TKDE' 22.

SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

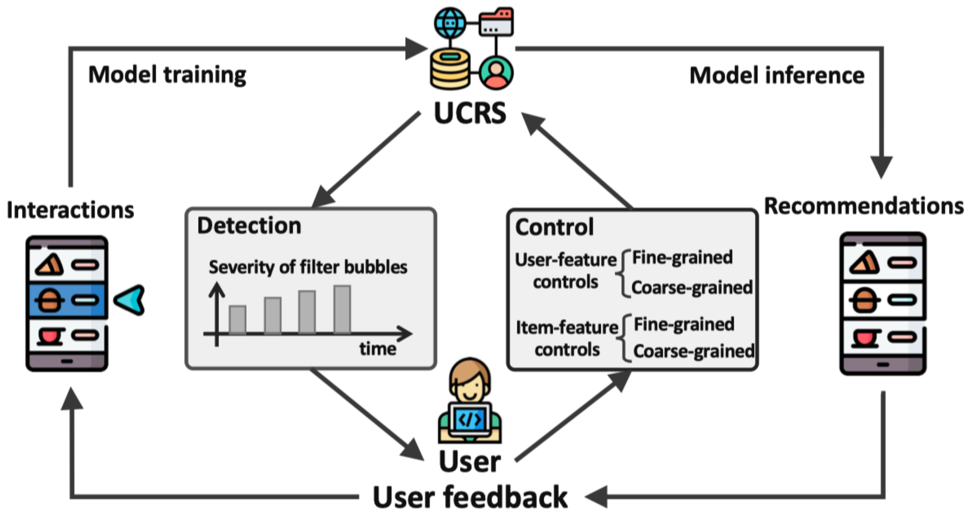
Counterfactual Recommendation

- Counterfactual for Alleviating Filter Bubbles
  - Filter bubbles in recommendation: RecSys emphasizes only a small set of items in the feedback loop.
  - Similar concepts: echo chamber, information cocoon.
  - Build causal models to interactive with users.

- Representative Work
  - Wang, et.al. User-controllable recommendation against filter bubbles. In SIGIR 2022.
  - Gao , et.al. CIRS: Bursting Filter Bubbles by Counterfactual Interactive Recommender System. In TOIS 2023.

Counterfactual for handling filter bubbles

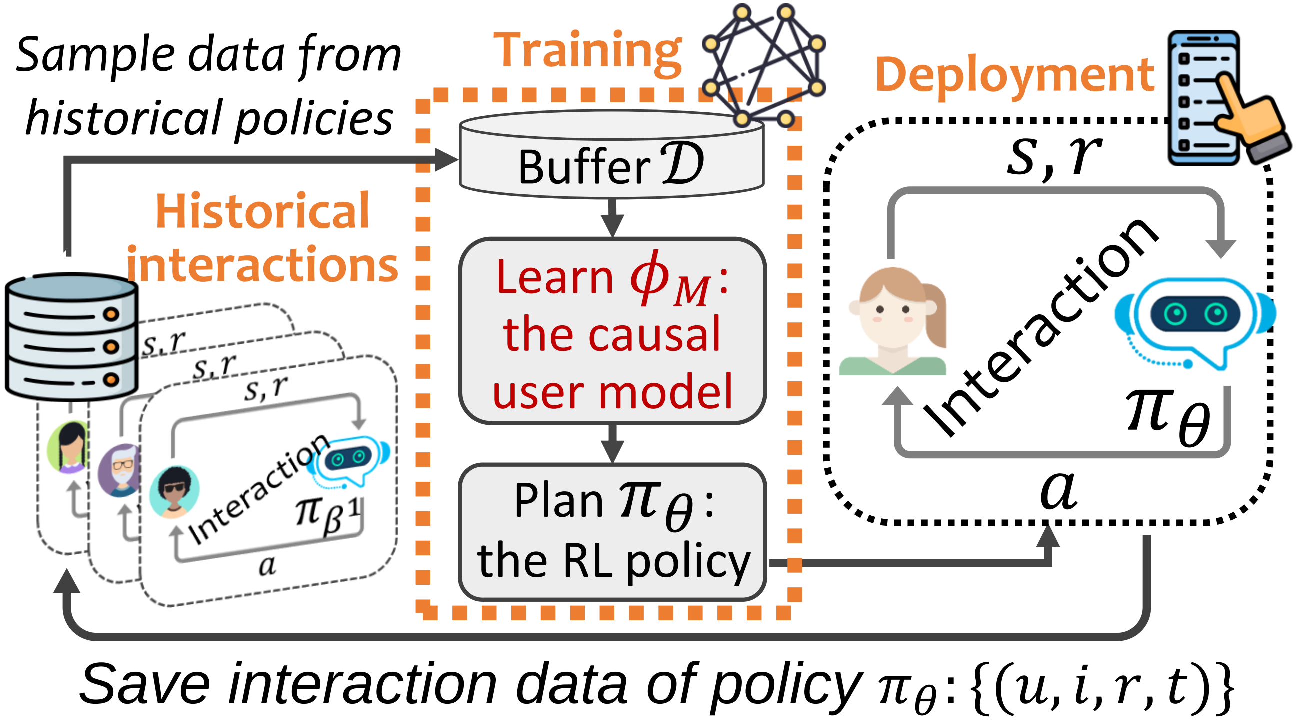
- **Filter bubbles** in recommendation: continually recommending many **homogeneous items**, isolating users from diverse contents.
- Solution: let users control the filter bubbles by directly adjusting recommendations.
- Two-level user controls regarding either a user or item feature.
  - Fine-grained level: increase the items *w.r.t.* a specified user or item feature.
    - For example, "more items liked by young users".
  - Coarse-grained level: no need to specify the target user/item group.
    - For example, "no bubble *w.r.t.* my age"
- A counterfactual imagination
  - Real-time response to user controls.
  - Need to reduce the effect of historical user representations.
  - **Counterfactual inference to mitigate the effect of out-of-date user interactions.**



Wang, et al. User-controllable recommendation against filter bubbles. In SIGIR 2022.

Counterfactual for handling filter bubbles

- Propose an unbiased **causal user model** ϕ_M in the model-based **offline reinforcement learning** (RL) framework to **disentangle** the intrinsic user interest from the **overexposure effect** of items.



Counterfactual IRS (CIRS) based on offline RL learning

- **Utilize counterfactual inference to disentangle and reduce the over-exposure effect on some items**

Gao , et.al. CIRS: Bursting Filter Bubbles by Counterfactual Interactive Recommender System. In TOIS 2023.

SCM for Recommendation

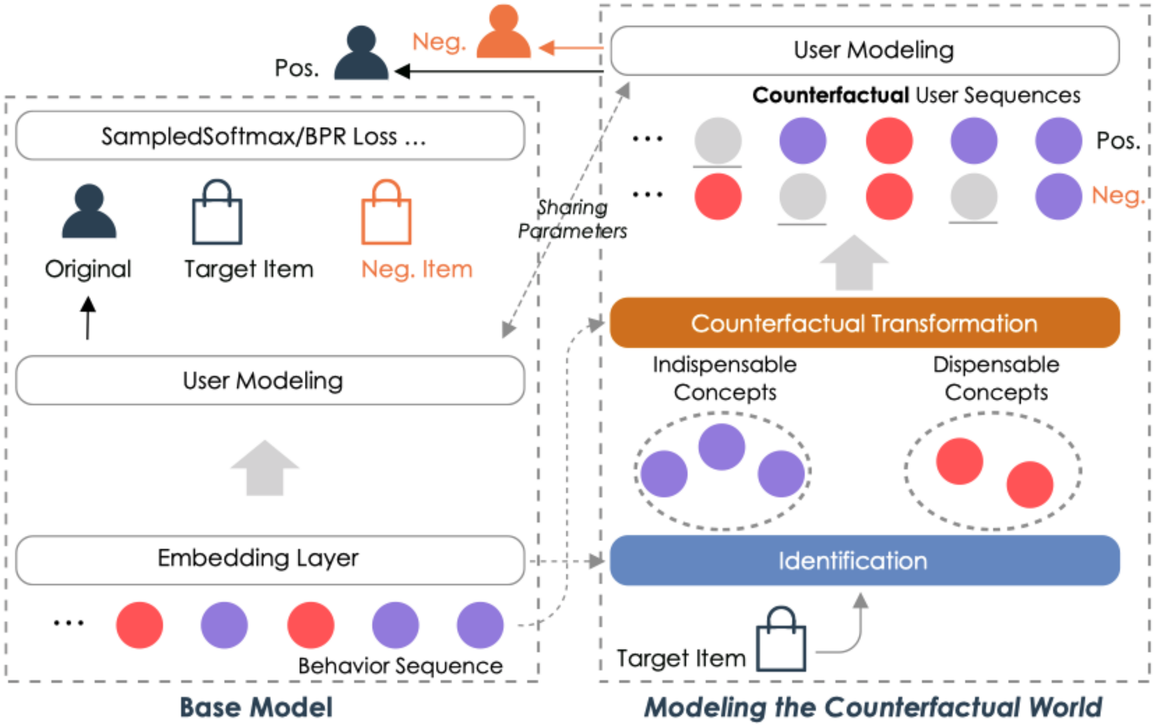
- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

Counterfactual Recommendation

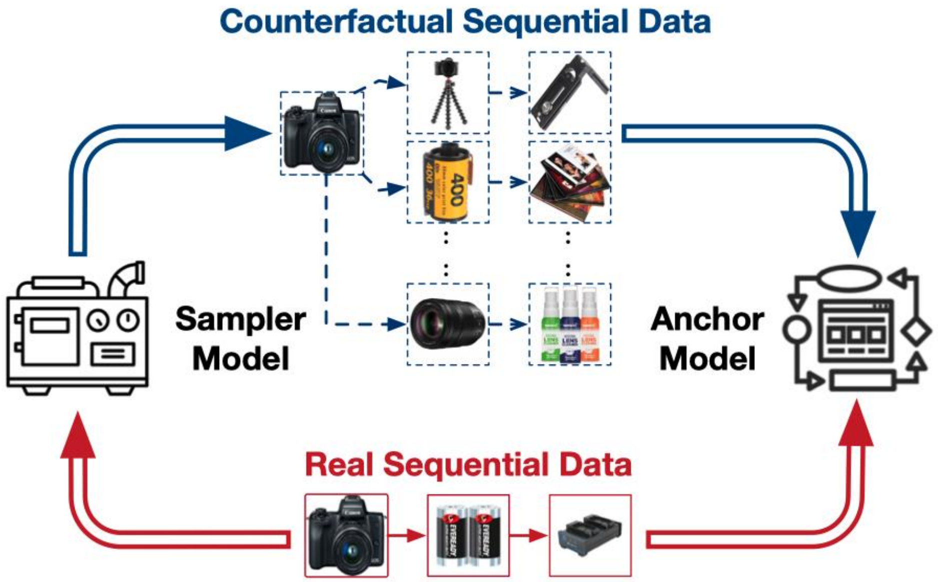
- Counterfactual data synthesis for alleviating data sparsity
  - Generate counterfactual interaction sequences for sequential recommendation.
  - Simulate the recommendation process and generate counterfactual samples, including recommendations and user feedback.
- Representative work
  - Zhang, et al. “Causerec: Counterfactual user sequence synthesis for sequential recommendation.” In SIGIR 2021.
  - Wang, et al. "Counterfactual data-augmented sequential recommendation." In SIGIR 2021.
  - Yang, Mengyue, et al. "Top-N Recommendation with Counterfactual User Preference Simulation." In CIKM 2021.

Counterfactual Data Synthesis

- Counterfactual data synthesis
  - Generate counterfactual interaction sequences for sequential recommendation.



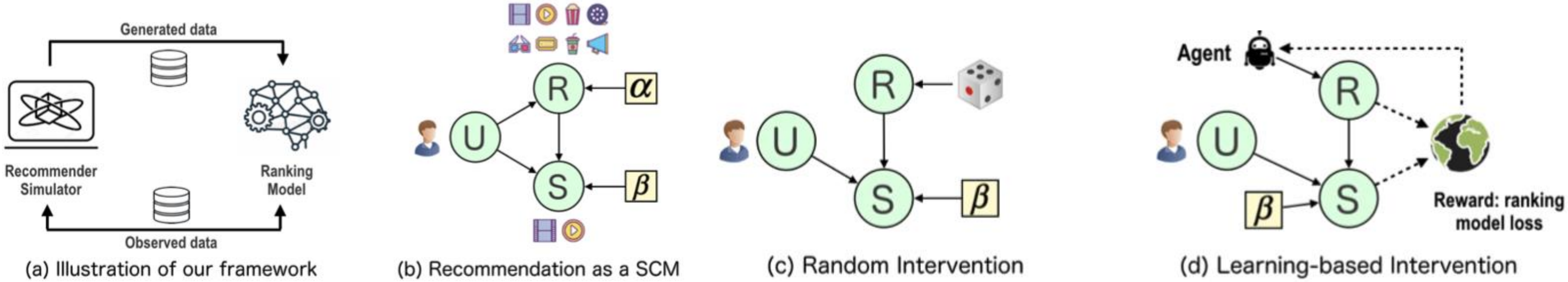
Zhang, et al. “Causerec: Counterfactual user sequence synthesis for sequential recommendation.” In SIGIR 2021.



Wang, et al. “Counterfactual data-augmented sequential recommendation.” In SIGIR 2021.

Counterfactual Data Synthesis

- Counterfactual data synthesis
  - Simulate the recommendation process and generate counterfactual samples, including recommendations and user feedback.
    1) Learn SCM from observed data to simulate the recommendation process.
    2) Conduct intervention on the recommendation list (R) to generate counterfactual samples.
    3) Use observed and generated data to train the ranking model.



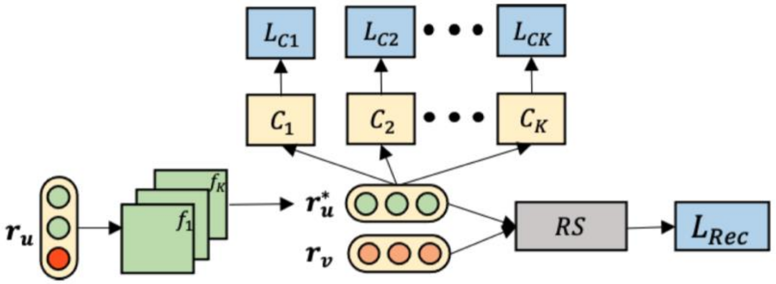
Yang, et al. “Top-N Recommendation with Counterfactual User Preference Simulation.” In CIKM 2021.

SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

Counterfactual Fairness

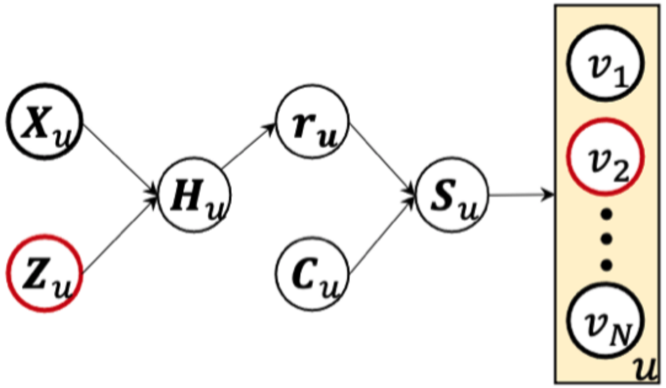
- Pursue fair recommendation for the users with different **sensitive attributes** (*e.g.*, age and gender).
- Counterfactual fair recommendation.
- Use **adversarial learning** to remove the sensitive information from user embedding (r_u).



DEFINITION 1 (COUNTERFACTUALLY FAIR RECOMMENDATION). *A recommender model is counterfactually fair if for any possible user u with features $\boldsymbol{X} = \boldsymbol{x}$ and $\boldsymbol{Z} = \boldsymbol{z}$:*

$$P\left(L_{\boldsymbol{z}} \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}\right) = P\left(L_{\boldsymbol{z}'} \mid \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}\right)$$

for all L and for any value $\boldsymbol{z}'$ attainable by $\boldsymbol{Z}$, where L denotes the Top-N recommendation list for user u.

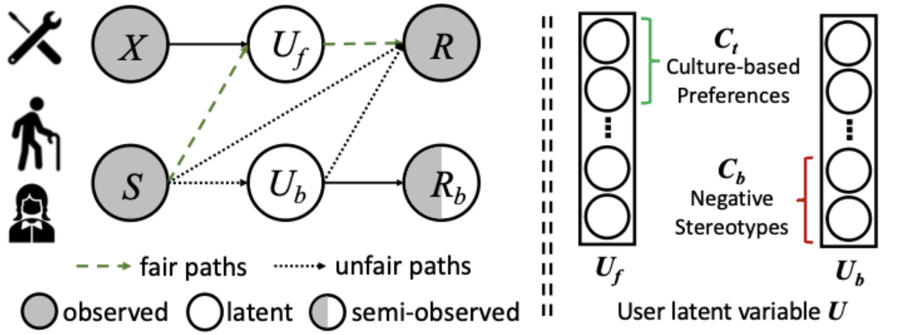


- X_u and Z_u are insensitive and sensitive features of the user u.
- H_u is the user interaction history.
- r_u is the user embedding.
- C_u is the candidate item set for u.
- S_u are the predicted scores over the candidate items.

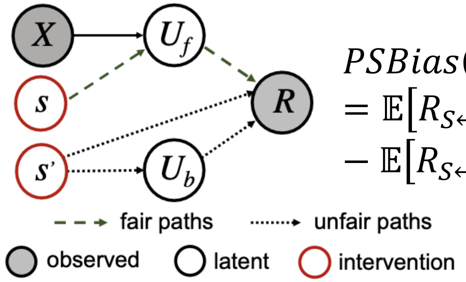
Li, et al. “Towards personalized fairness based on causal notion.” In SIGIR 2021.

Counterfactual Fairness

- **Path-specific (PS) counterfactual fairness**
  - PS fair recommendation
    - **eliminate the unfair influences** of sensitive features (*e.g.*, race)
    - **preserve fair influences** of sensitive features (*e.g.*, chopsticks for East-Asian users).
  - Calculate and remove **PS bias** based on **path-specific counterfactual inference.**



X: non-sensitive user features S: sensitive user features
U_f: user fair latent variable U_b: user bias latent variable
R: observed ratings R_b: semi-observed unfair ratings



$$PSBias(\mathrm{x}, \mathrm{s}, \mathrm{s}') = \mathbb{E}\left[R_{S\leftarrow s'}\left(U_{f,S\leftarrow s}, U_{b,S\leftarrow s'}\right) \middle| \mathrm{X}=\mathrm{x}, \mathrm{S}=\mathrm{s}\right] - \mathbb{E}\left[R_{S\leftarrow s}\left(U_{f,S\leftarrow s}, U_{b,S\leftarrow s}\right) \middle| \mathrm{X}=\mathrm{x}, \mathrm{S}=\mathrm{s}\right]$$

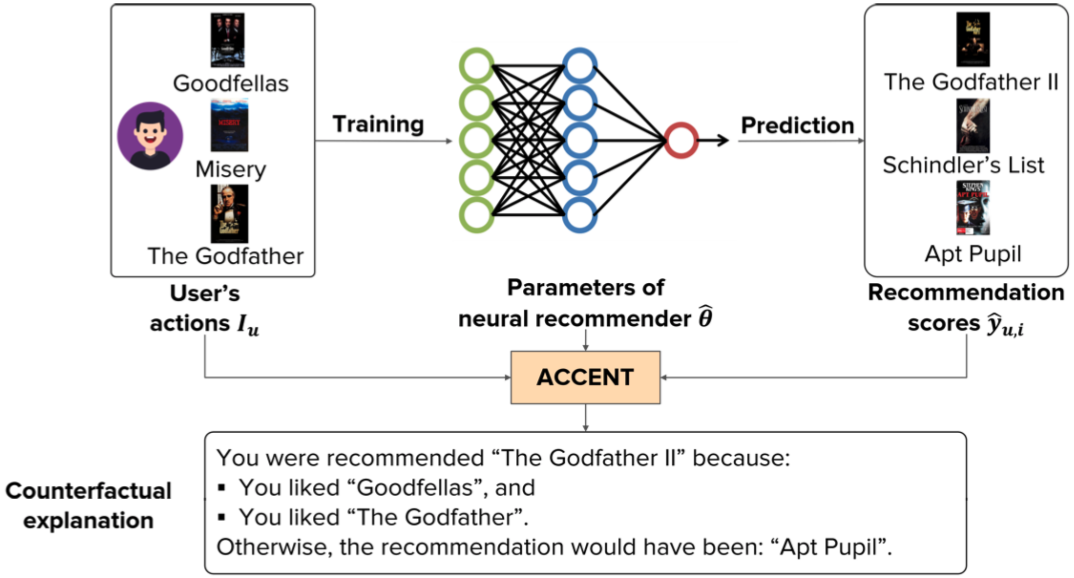
Zhu, et al. "Path-Specific Counterfactual Fairness for Recommender Systems." In KDD 2023.

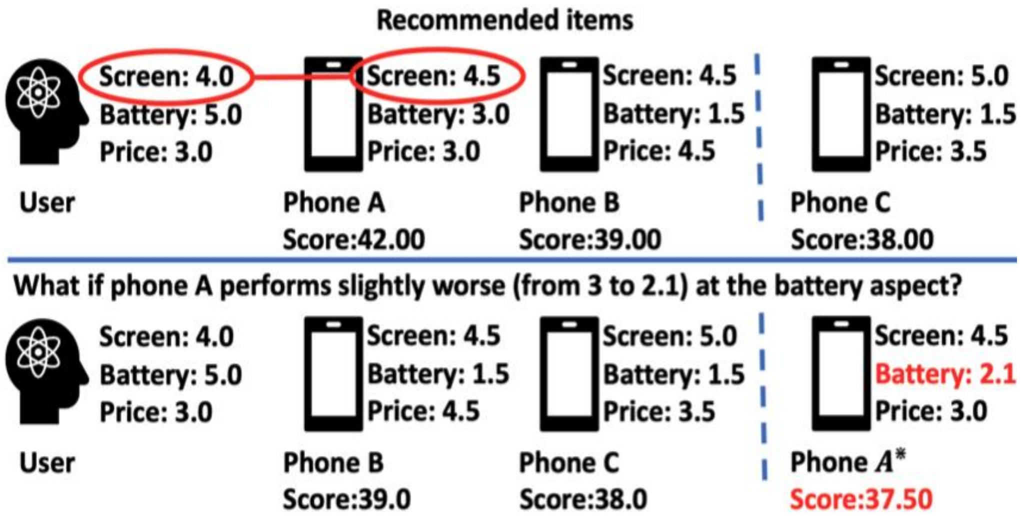
SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

Counterfactual Explanation

- Generate explanation by counterfactual thinking.
- Find the minimal changes that lead to a different recommendation.
- Identify the most critical features causing the recommendations.





Tran, et al. "Counterfactual Explanations for Neural Recommenders." In SIGIR 2021.

Tan, et al. "Counterfactual explainable recommendation." In CIKM 2021.

SCM for Recommendation

- Dealing with confounding structures in recommendation (Yang Zhang)
  - Confounding in recommendation.
  - Deal with observed confounders.
  - Deal with unobserved confounders.
- Considering colliding structures in recommendation (Yang Zhang)
  - Colliders in recommendation
  - Modeling the colliding effect
- Counterfactual recommendation (Wenjie Wang)
  - Counterfactual inference for debiasing
  - Counterfactual inference against filter bubbles
  - Counterfactual data synthesizing
  - Counterfactual fairness
  - Counterfactual explanation
  - Causal modeling for OOD generalization

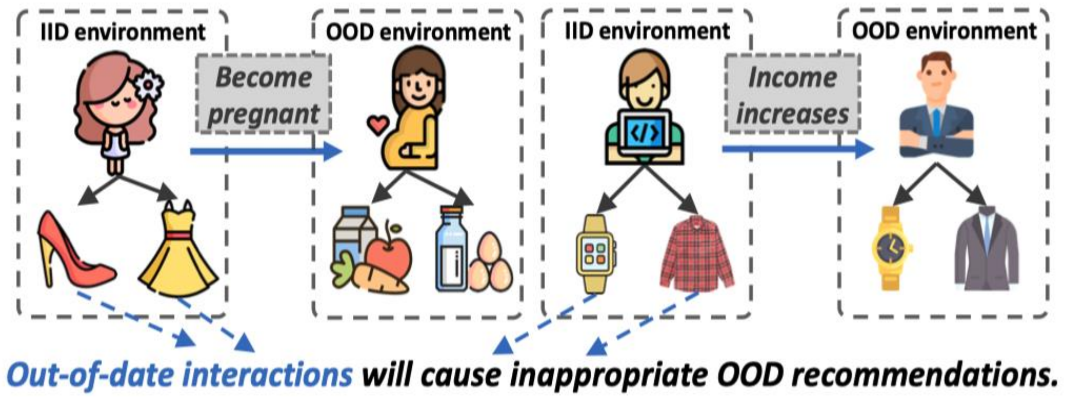
Counterfactual Recommendation

- Causal Modeling for OOD Recommendation
  - The interaction distribution is shifting over time in recommendation.
  - Leverage causal modeling to enhance the recommender generalization.

- Representative Work
  - Wang et.al. Causal representation learning for out-of-distribution recommendation. In WWW 2022.
  - He et al. CausPref: Causal Preference Learning for Out-of-Distribution Recommendation. In WWW 2022.
  - Wang et al. Causal Disentangled Recommendation Against User Preference Shifts. In TOIS 2023.
  - Zhang et al. Invariant Collaborative Filtering to Popularity Distribution Shift. In WWW 2023.

Causal Modeling for OOD Recommendation

- **User preference is shifting over time.**
- **Reason** of the preference shifts: **change of user features.**
  - User features → preference → interactions.
- Explore OOD recommendation under two settings:
  - OOD recommendation with **observed user features**. (*e.g.,* increased consumption levels and changed location)
  - OOD recommendation with **unobserved user features.** (*e.g.,* friend recommendations, hot event, and context factors)



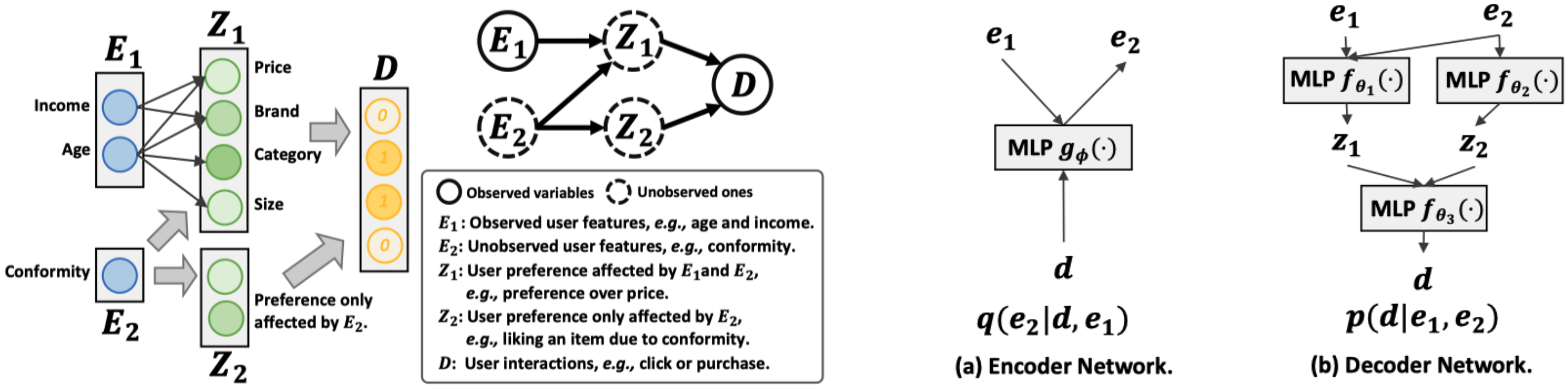
Wang, et al. "Causal Representation Learning for Out-of-Distribution Recommendation." In *WWW 2022.*

Causal Modeling for OOD Recommendation

- **OOD recommendation with observed user features.**

1) Figure out the mechanism how feature shifts affect user preference.
   - User features → preference → interactions.
   - Leverage VAE framework to **model the causal relations** behind the interaction generation process.

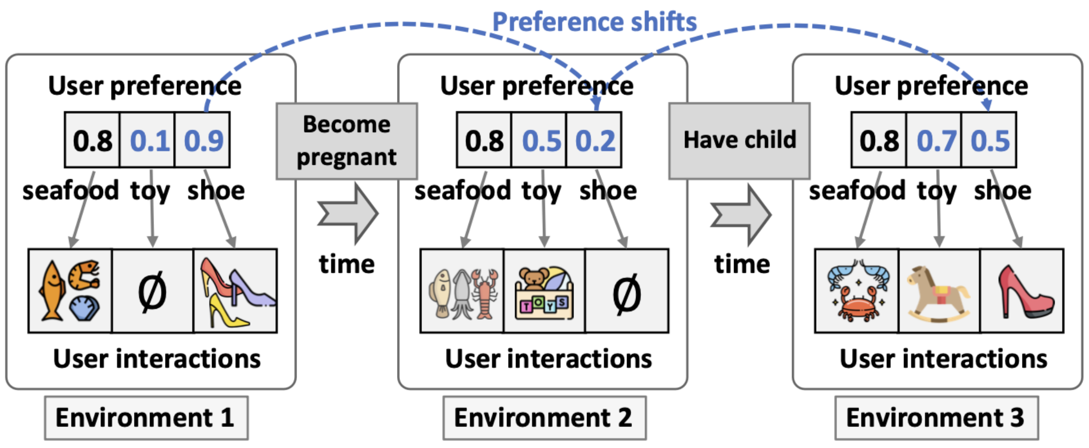
2) Mitigate the effect of out-of-date interactions.
   - **Counterfactual inference**: what the user preference would be if the out-of-date interactions were removed?

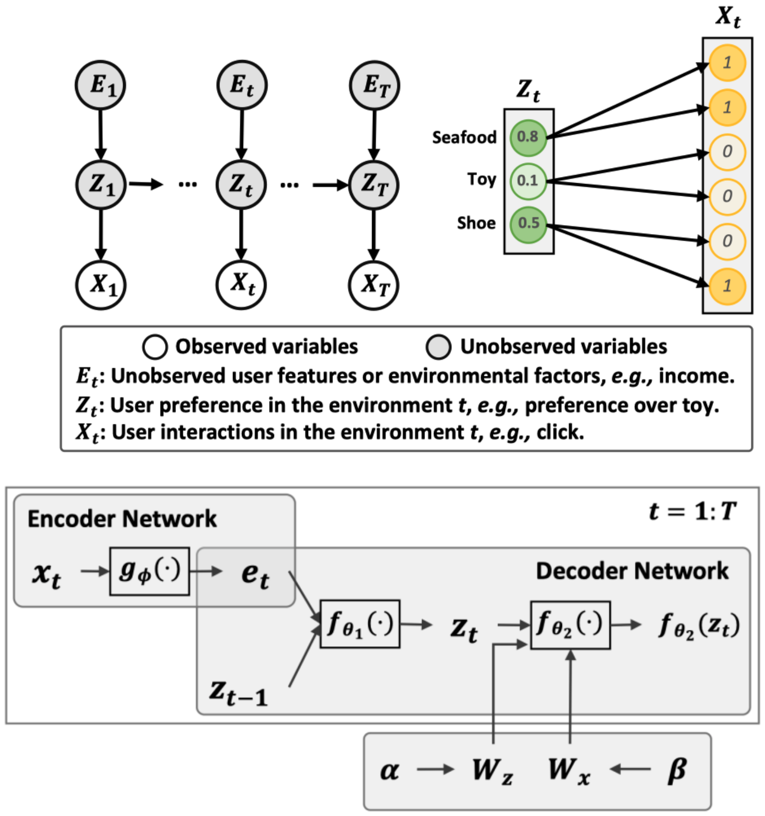


Wang, et al. "Causal Representation Learning for Out-of-Distribution Recommendation." In *WWW 2022.*

Causal Modeling for OOD Recommendation

- **OOD recommendation with unobserved user features.**
  - Unobserved factors cause preference shifts.
  - Example: friend recommendations, hot event, and other environmental factors.





Wang, et al. "Causal Disentangled Recommendation Against User Preference Shifts." In *TOIS 2023.*

Papers on Counterfactual Recommendation

- Wang, et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue. In SIGIR 2021.
- Wei, et al. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system. In KDD 2021.
- Zihao Zhao et al. Popularity Bias Is Not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. In TKDE (2022).
- Gang Chen et al. Unbiased Knowledge Distillation for Recommendation. In WSDM 2023.
- Wang, et.al. User-controllable recommendation against filter bubbles. In SIGIR 2022.
- Gao , et.al. CIRS: Bursting Filter Bubbles by Counterfactual Interactive Recommender System. In TOIS 2023.
- Zhang, et al. “Causerec: Counterfactual user sequence synthesis for sequential recommendation.” In SIGIR 2021.
- Wang, et al. "Counterfactual data-augmented sequential recommendation." In SIGIR 2021.
- Yang, Mengyue, et al. "Top-N Recommendation with Counterfactual User Preference Simulation." In CIKM 2021.
- Li, et al. “Towards personalized fairness based on causal notion.” In SIGIR 2021.
- Yaochen Zhu et. al. Path-Specific Counterfactual Fairness for Recommender Systems. In KDD 2023.
- Tran, et al. “Counterfactual Explanations for Neural Recommenders.” In SIGIR 2021.
- Tan, et al. “Counterfactual explainable recommendation.” In CIKM 2021.
- Wang, et.al. Causal representation learning for out-of-distribution recommendation. In WWW 2022.
- He et al. CausPref: Causal Preference Learning for Out-of-Distribution Recommendation. In WWW 2022.
- Wang et al. Causal Disentangled Recommendation Against User Preference Shifts. In TOIS 2023.
- Zhang et al. Invariant Collaborative Filtering to Popularity Distribution Shift. In WWW 2023.

Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

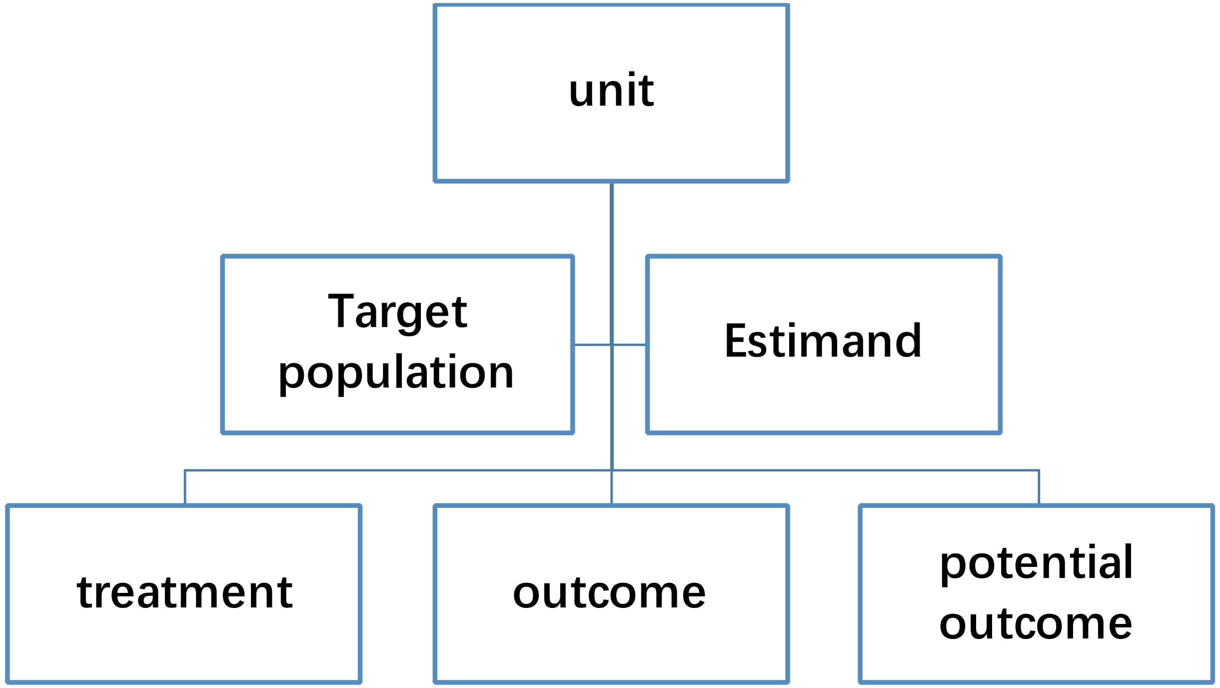
Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
  - Bias-Variance Trade-Off
  - Robust to Small Propensities (Data Sparsity)
  - Robust to Pseudo-Labelings
  - Mitigating/Eliminating Unmeasured Confounding
  - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Potential Outcome Framework



Key elements in Potential Outcome (PO) framework

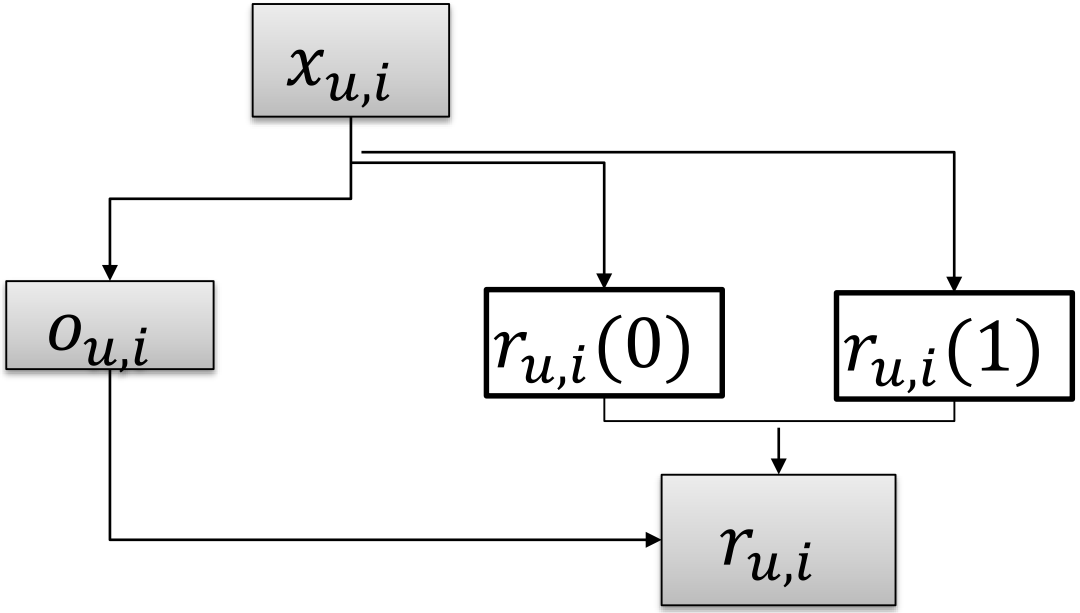
- **Unit**: the most fine-grained research subject.
- **Target population**: the population that we want to make an inference/prediction on.
- **Causal estimand**: the causal parameter, providing a recipe for answering the scientific question of interest from any hypothetical data whenever it is available.

Imbens, G. W. and D. B. Rubin (2015). "Causal Inference For Statistics Social and Biomedical Science", Cambridge University Press.

Example: PO Framework in RS

- *Unit*: a user-item pair (u, i).
- *Target population*: the set of all user-item pairs $\mathcal{D} = \mathcal{U} \times \mathcal{I}$.
- *Feature*: the feature $x_{u,i}$ describes user-item pair (u, i).
- *Treatment*: $o_{u,i} \in \{1, 0\}$. It is the exposure status of (u, i), where $o_{u,i} = 1$ or 0 denotes item i is exposed to user u or not.
- *Outcome*: the feedback $r_{u,i}$ of user-item pair (u, i).
- *Potential outcome*: $r_{u,i}(o)$ for $o \in \{0, 1\}$. It is the outcome that would be observed if $o_{u,i}$ had been set to o.

Example: PO Framework in RS



In RS, we usually want to answer the intervention question "if recommending an item to a user, what would the feedback be". Formally, the causal estimand is

$$\mathbb{E}(r_{u,i}(1) \mid x_{u,i}), \tag{1}$$

which requires to predict the potential outcome $r_{u,i}(1)$ using feature $x_{u,i}$.

79

Example: PO Framework in RS

Example 1: video websites.

- r_{ui}: the true rating of user u for video i.
- o_{ui}: observing indictor.
  $o_{ui} = 1 \iff r_{ui}$ is observed

Table 1: Data structure of example 1.

o_{ui}	x_{ui}	$r_{ui}(1)$
1	✓	✓
1	✓	✓
1	✓	✓
0	✓	
0	✓	
0	✓	

We can regard the observing indicator $o_{u,i}$ as the treatment, and define $r_{u,i}(1)$ as the true potential rating if $o_{u,i} = 1$ for all user-item pairs. Here we use $r_{u,i}(1)$ instead of $r_{u,i}$ is to underline that the potential outcomes of interest are partially observable.

Goal: predict the potential outcome $r_{u,i}(1)$ using feature $x_{u,i}$.

Example: PO Framework in RS

Example 2: advertising CTR Predication.

- r_{ui}: $r_{ui} = 1$ if u clicks on item i; $r_{ui} = 0$ otherwise.
- o_{ui}: $o_{ui} = 1$ if item i is exposed to u; $o_{ui} = 0$ otherwise.
- CTR: $\mathbb{E}[r_{ui}(1)|x_{u,i}] = \mathbb{P}(r_{ui}(1) = 1|x_{u,i})$.

o_{ui}	x_{ui}	r_{ui}	$r_{ui}(1)$
1	✓	✓	✓
1	✓	✓	✓
1	✓	✓	✓
0	✓	✓	
0	✓	✓	
0	✓	✓	

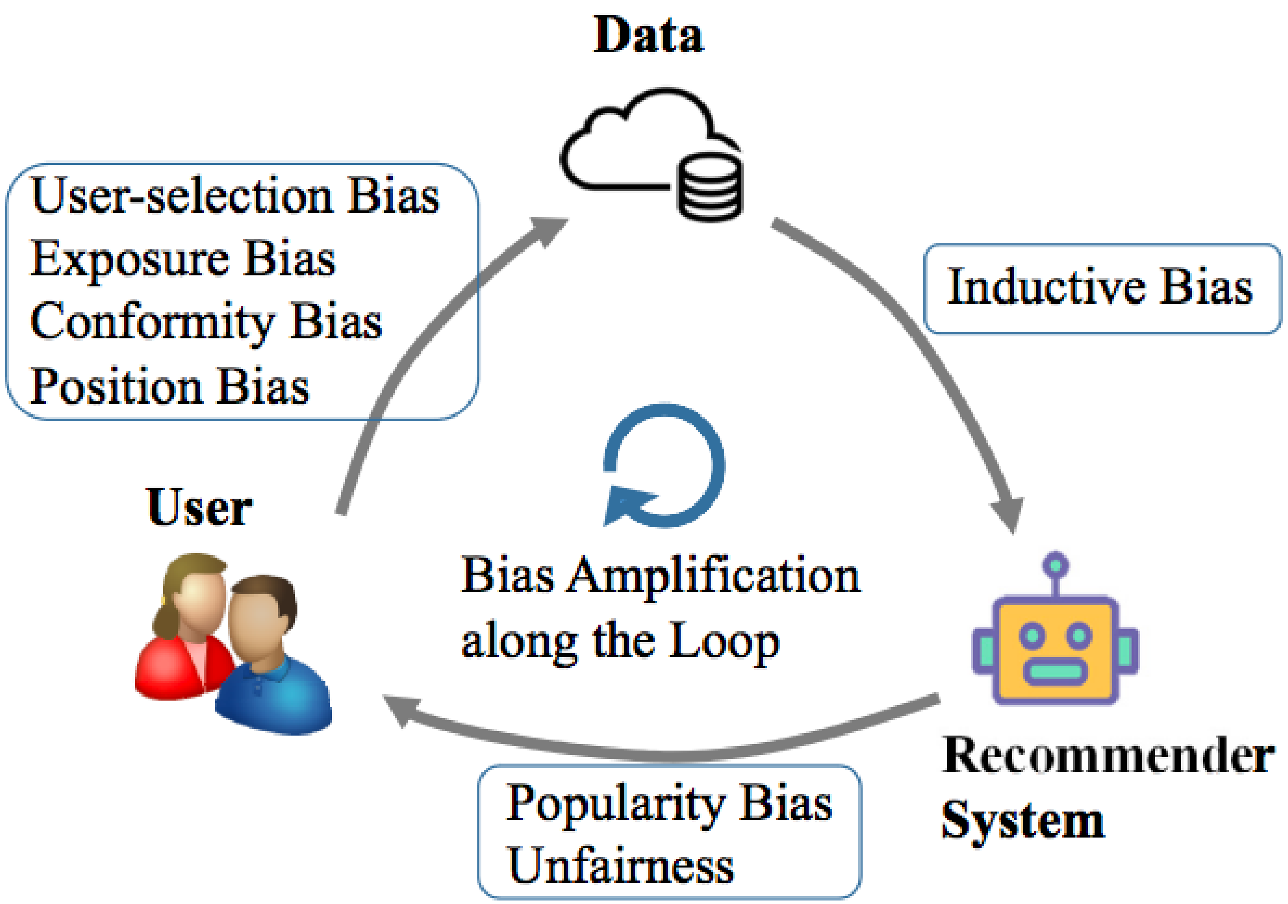
Example 3: advertising post-click CVR Predication.

- r_{ui}: $r_{ui} = 1$ if user u purchases item i; $r_{ui} = 0$ otherwise.
- o_{ui}: $o_{ui} = 1$ if user u clicks item i $o_{ui} = 0$ otherwise.
- post-click CVR: $\mathbb{E}[r_{ui}(1)|x_{u,i}] = \mathbb{P}(r_{ui}(1) = 1|x_{u,i})$.

PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
  - Bias-Variance Trade-Off
  - Robust to Small Propensities (Data Sparsity)
  - Robust to Pseudo-Labelings
  - Mitigating/Eliminating Unmeasured Confounding
  - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Biases in RS



Chen et al. TOIS 2023. Bias and Debias in Recommender System: A Survey and Future Directions.

Bias is usually Evil

- Economic
  - Bias affects recommendation accuracy
  - Bias hurts user experience, causing the losses of users
  - Unfairness incurs the losses of item providers

- Society
  - Bias can reinforce discrimination of certain user's groups
  - Bias decreases the diversity and intensify the homogenization of users

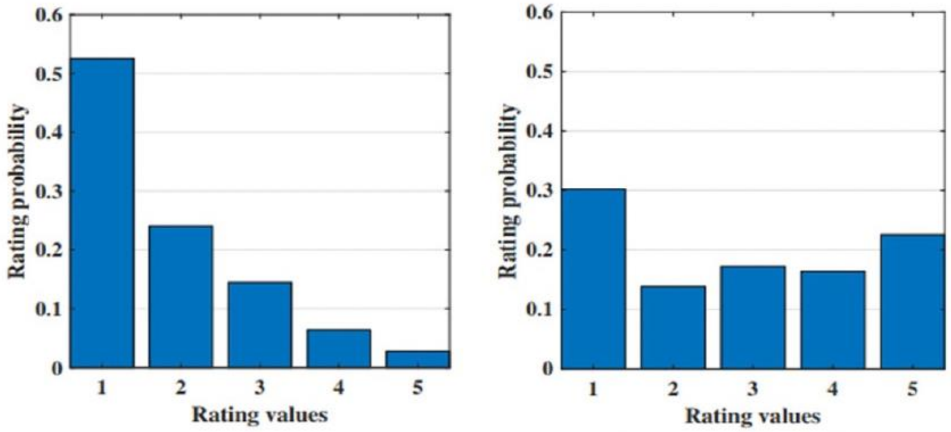
Selection Bias

- Definition: *Selection bias happens in explicit feedback data as users are free to choose which items to rate, so that the observed ratings are not a representative sample of all ratings.*

3	4	2	5
1	3	2	5
2	3	4	4

Selection bias →

3	4		5
	3		5
	3	4	4

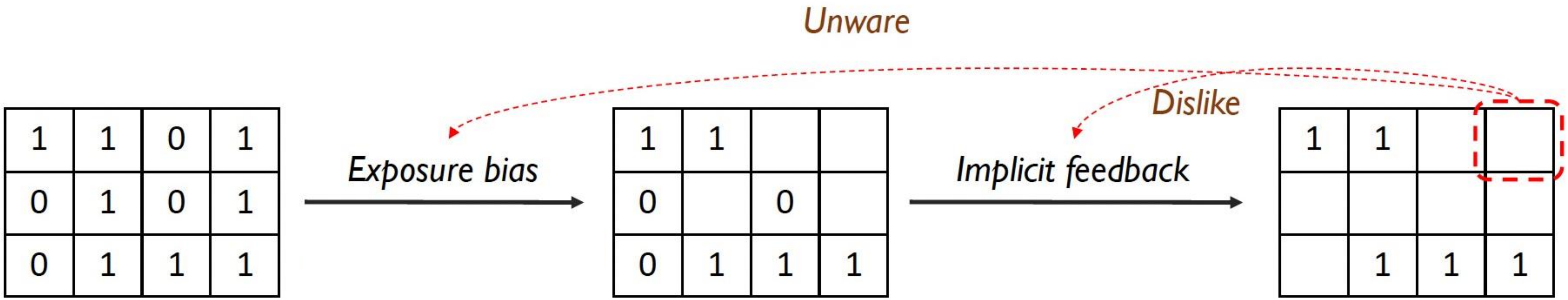


(a) Random (b) User-selected

[1] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In ICML.
[2] B. M. Marlin, R. S. Zemel, S. Roweis, and M. Slaney, "Collaborative filtering and the missing at random assumption," in UAI, 2007

Exposure Bias

- Definition: *Exposure bias happens in implicit feedback data as users are only exposed to a part of specific items.*



Conformity Bias

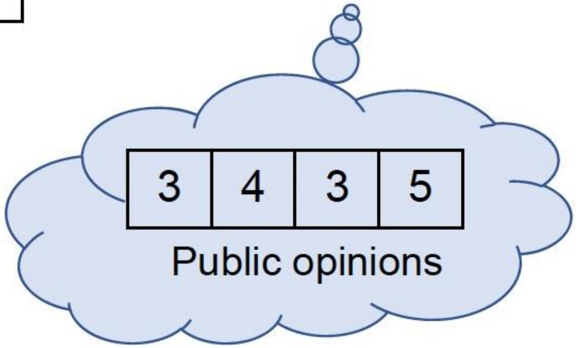
- Definition: Conformity bias *happens as users tend to behave similarly to the others in a group, even if doing so goes against their own judgment.*

3	4		5
	3		4
	3	4	3

Conformity bias →

$$p_T(r|u,i) \neq p_D(r\,|\,u,i)$$

3	4		5
	3		5
	3	3	4



Bias Formalization

Lack of formal definitions of various biases under PO framework in RS!

Biases in RS

Selection bias
Conformity Bias
Exposure Bias
Position Bias
Inductive Bias
Popularity Bias
……

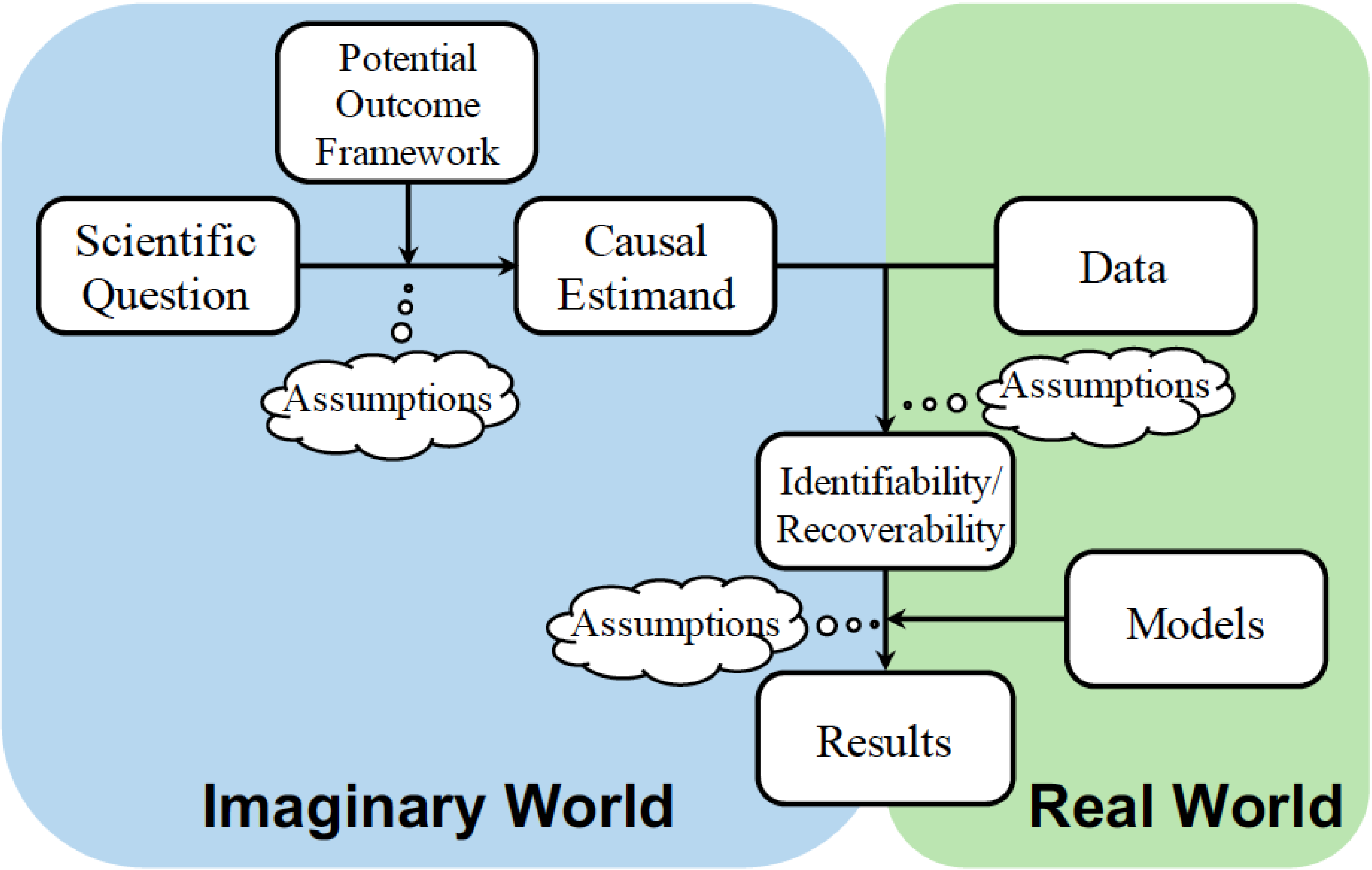
Biases in Causal Inference

Noncompliance
Interference bias
Unmeasured confounding
Confounding bias
Selection bias
Model assumption

……

Peng Wu*, Haoxuan Li*, Yuhao Deng, Wenjie Hu, Quanyu Dai, Zhenhua Dong, Jie Sun, Rui Zhang, Xiao-Hua Zhou (2022), ''On the Opportunity of Causal Learning in Recommendation Systems: Foundation, Estimation, Prediction and Challenges'', IJCAI 22.

Biases in Causal Inference



We need a variety of assumptions to climb from association (data) to causality (causal conclusions), violating these assumptions may result in various biases.

Bias Formalization under PO Framework

	Assumptions	Biases in causal inference	Biases in [Chen *et al.*, 2020]
Define causal estimands	SUTVA(a) SUTVA(b)	undefined interference bias	position bias conformity bias
Recoverability	consistency positivity exchangeability conditional exchangeability random sampling	noncompliance undefined confounding bias hidden confounding bias selection bias	undefined exposure bias popularity bias undefined user/model selection bias, exposure bias
Model	model specification	model mis-specification	inductive bias

Table 1: New perspective of biases in RS.

- We can define the descriptive biases in RS formally using the rigorous syntax of causal inference.
- It also provides an opportunity to apply the existing causal inference methods to RS.
- In addition, for the unique characteristics of RS, we expect that a series of new methods will be developed by weakening or substituting the assumptions.

Peng Wu*, Haoxuan Li*, Yuhao Deng, Wenjie Hu, Quanyu Dai, Zhenhua Dong, Jie Sun, Rui Zhang, Xiao-Hua Zhou (2022), ''On the Opportunity of Causal Learning in Recommendation Systems: Foundation, Estimation, Prediction and Challenges'', IJCAI 22.

PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
  - Bias-Variance Trade-Off
  - Robust to Small Propensities (Data Sparsity)
  - Robust to Pseudo-Labelings
  - Mitigating/Eliminating Unmeasured Confounding
  - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Selection Bias Formulation

Formalization of "what would the response be if recommending an item to a user?":

- **Unit:** a user-item pair (u, i);
- **Target population:** the set of all user-item pairs: $\mathcal{D} = \mathcal{U} \times I$;
- **Feature:** the feature $x_{u,i}$ describes user-item pair (u, i);
- **Treatment:** $o_{u,i} \in \{0,1\}$, which is the **exposure** indicator of (u, i);
- **Observed Outcome:** the response $r_{u,i}$ of (u, i), e.g., **watch**;
- **Potential outcome:** $r_{u,i}(o)$ for $o \in \{0,1\}$, which is the outcome that would be observed if $o_{u,i}$ had been set to o.

o_{ui}	x_{ui}	r_{ui}	$r_{ui}(1)$
1	✓	✓	✓
1	✓	✓	✓
1	✓	✓	✓
0	✓	✓	
0	✓	✓	
0	✓	✓	

$$\mathrm{P}(r_{u,i} = 1 | X_{u,i} = x_{u,i}, o_{u,i} = 1) \Rightarrow \mathrm{P}(r_{u,i}(1) = 1 | X_{u,i} = x_{u,i})$$

Associational Definition → **Causal Estimand**

Ideal Prediction Loss in RS

Let f_ϕ be a recommender model used to predict $r_{u,i}(1)$.

Ideal Loss: If all potential outcomes $\{r_{u,i}(1) : (u, i) \in \mathcal{D}\}$ were observed, the ideal loss function for training ϕ is

$$\mathcal{L}_{ideal}(\phi) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e_{u,i}, \tag{2}$$

where $e_{u,i} = L(r_{ui}(1), f_\phi(x_{u,i}))$ is the prediction error, such as the least square loss:

$$e_{u,i} = (f_\phi(x_{u,i}) - r_{u,i}(1))^2. \tag{3}$$

Noticing that $e_{u,i}$ is computable only when $o_{u,i} = 1$, $L_{ideal}(\phi)$ is infeasible. As such, our target is constructing estimators that approximate to $L_{ideal}(\phi)$.

Debiasing Strategies: Overview

- Re-weighting
  - Giving weights for each instance to re-scale their contributions on model training
- Re-labeling
  - Giving a new pseudo-label for the missing or biased data
- Generative Modeling
  - Assuming the generation process of data and reduces the biases accordingly

Propensity Score for Biases (Reweighting)

3	4	2	5
1	3	2	5
2	3	4	4

Selection bias

$$p_T(u,i) \neq p_D(u,i)$$

3	4		5
	3		
2	3	4	4

Reweighting

3	4		5
	3		
2	3	4	4

$$L_{IPW}(S,q) = \sum_{x \in \pi_q} \Delta_{IPW}(x, y \mid \pi_q)$$

$$= \sum_{x \in \pi_q, o_q^x = 1, y = 1} \frac{\Delta(x, y \mid \pi_q)}{\boxed{P(o_q^x = 1 \mid \pi_q)}}$$

Simple and straightforward.
Theoretical soundness.

High Variance.
Difficult to set proper propensity score.
Requires positivity.

Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In ICML

T. Joachims, A. Swaminathan, and T. Schnabel, "Unbiased learning-to-rank with biased feedback," in WSDM, 2017, pp. 781–789

Data Imputation (Relabeling)

True Preference

3	4	2	5
1	3	2	5
2	3	4	4

$\xrightarrow{\text{Selection bias}}$ $p_T(u,i) \neq p_D(u,i)$

Training data

3	4		5
	3		5
	3	4	4

Data imputation $\dashrightarrow$

Imputation data

3	4	2	5
2	3	2	5
2	3	4	4

- Relabeling: assigns pseudo-labels for missing data.

$$\arg\min_{\theta} \sum_{u,i} \hat{\delta}\left(\boxed{r_{ui}^{o\&i}}, f(u,i \mid \theta)\right) + \mathrm{Reg}(\theta)$$

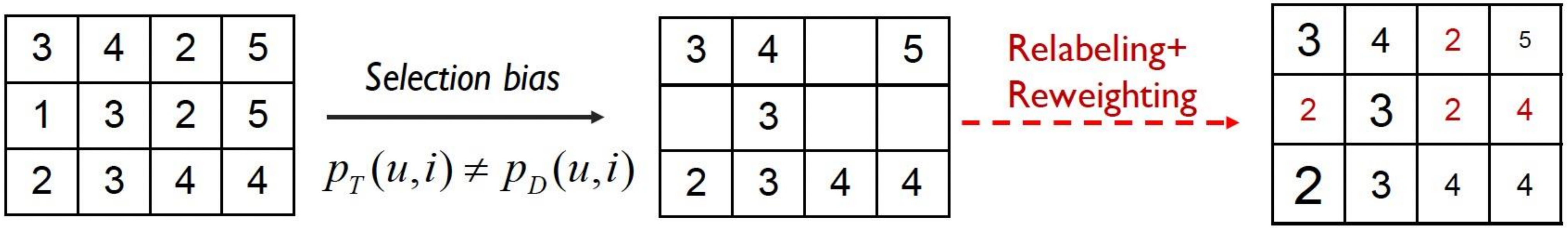
Simple and straightforward.

Sensitive to the imputation strategy. Imputing proper pseudo-labels is more difficult.

H. Steck, "Training and testing of recommender systems on data missing not at random," in KDD, 2010, pp. 713–722.

X. Wang, R. Zhang, Y. Sun, and J. Qi, "Doubly robust joint learning for recommendation on data missing not at random," in ICML, 2019, pp. 6638–6647

Doubly Robust (Relabeling+Reweighting)



- Doubly Robust: combines IPS and data imputation for robustness.

$$\hat{L}_{DR} = \underbrace{\sum_{(u,i)\in D_T} \frac{1}{\rho_{ui}}}_{\text{IPS}} \left(\delta\left(\hat{r}_{ui}, r_{ui}\right)\right) + \sum_{u\in U, i\in I} \underbrace{\left(1-\frac{O_{ui}}{\rho_{ui}}\right)\delta\left(\hat{r}_{ui}, m_{ui}\right)}_{\text{Imputation}}$$

$$O_{ui} = \mathbf{I}[(u,i) \in D_T]$$

Low Variance.

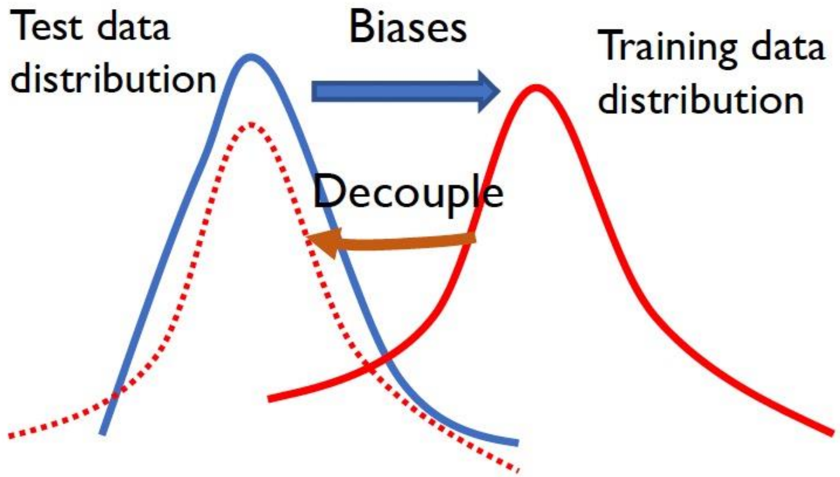
Relatively robust to the propensity score and imputation value.

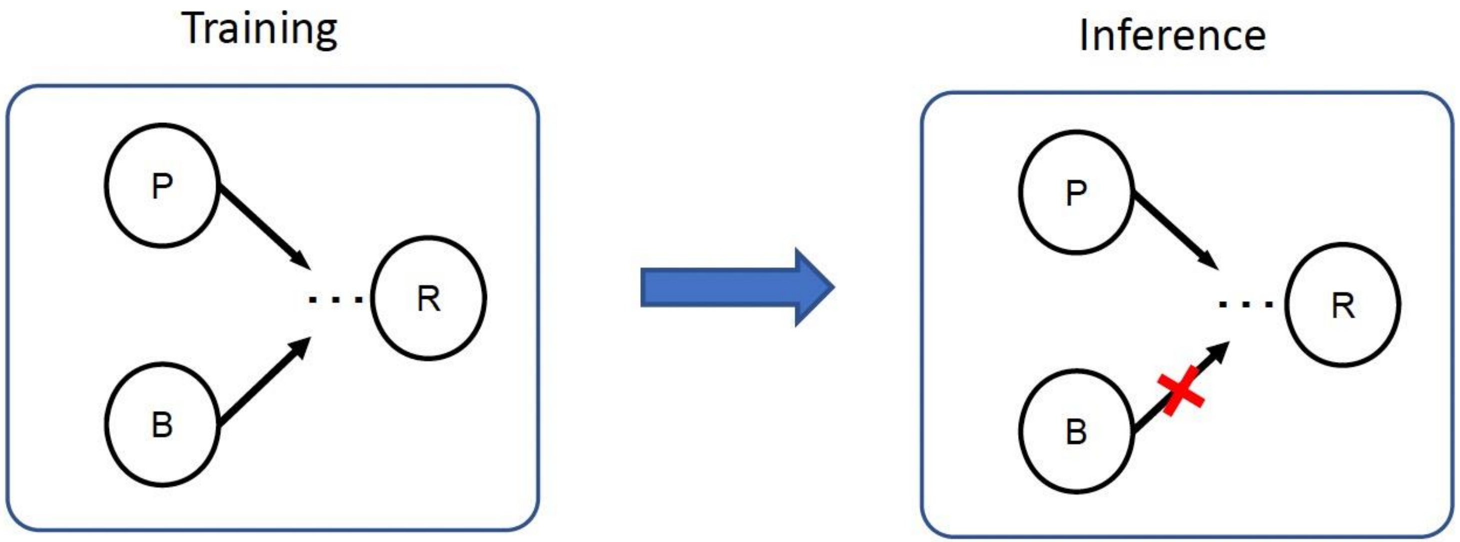
Requires proper imputation or propensity strategy.

Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2019. Doubly robust joint learning for recommendation on data missing not at random. In ICML.

Generative Modeling

- Basic idea: assuming the generation process of data to decouple the effect of user true preference from the bias.





PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
  - Bias-Variance Trade-Off
  - Robust to Small Propensities (Data Sparsity)
  - Robust to Pseudo-Labelings
  - Mitigating/Eliminating Unmeasured Confounding
  - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Basic Methods

EIB: Error Imputation Based Estimator

- Try to recover the whole data space
- Impute the prediction error of unobserved data

IPS: Inverse Propensity Score Estimator

- Model the missing mechanism to obtain propensity
- Adjust the distribution of observed data through reweighting

DR: Doubly Robust Learning Estimator

- Double robustness: unifies the advantages of EIB and IPS

DR methods: DR, DR-JL, Multi-DR, MRDR

- The generalization bound of DR methods

$$\mathcal{L}_{EIB}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} [o_{u,i} e_{u,i} + (1 - o_{u,i}) \hat{e}_{u,i}]$$

$$\mathcal{L}_{IPS}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}}$$

$$\mathcal{L}_{DR}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left[\frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}} + (1 - \frac{o_{u,i}}{\hat{p}_{u,i}}) \hat{e}_{u,i} \right]$$

$$\mathcal{L}_{\mathrm{e}}(\theta, \phi) = \sum_{u,i\in\mathcal{O}} \frac{(\hat{e}_{u,i} - e_{u,i})^2}{\hat{p}_{u,i}}$$

$$\mathcal{L}_{ideal}(\hat{\mathbf{R}}^*) \le \underbrace{\boxed{\mathcal{L}_{DR}(\hat{\mathbf{R}}^*)}}_{\text{Error Term}} + \underbrace{\boxed{\frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{|p_{u,i} - \hat{p}_{u,i}|}{\hat{p}_{u,i}} |\hat{e}_{u,i} - e^*_{u,i}|}}_{\text{Bias Term}} + \underbrace{\boxed{\sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} (\frac{\hat{e}_{u,i} - e^*_{u,i}}{\hat{p}_{u,i}})^2}}}_{\text{Variance Term}}$$

Based on the theoretic analysis, the generalization of existing DR methods could still be improved by better controlling the upper bound!

Existing Methods	Weakness
EIB	High bias
IPS, Multi-IPW	High variance
DR, DR-JL, Multi-DR	Still suffer bias and variance
MRDR	Still suffer from bias

Limitations of IPS and DR methods

TABLE I: Comparison of various debiasing estimators.

	IPS	SNIPS	EIB	DR	TDR
Doubly robust	$\times$	$\times$	$\times$	$\checkmark$	$\checkmark$
Low variance	$\times$	$\times$	$\checkmark$	o	$\checkmark$
Robust to small propensities	$\times$	o	$\checkmark$	$\times$	$\checkmark$
Without extrapolation	$\checkmark$	$\checkmark$	$\times$	o	o
Boundedness	$\times$	$\checkmark$	$\checkmark$	$\times$	$\checkmark$

Note: symbols $\checkmark$, o and $\times$ denote good, medium and bad, respectively.

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, “TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations.” ICLR 23.

Five Desired Properties

- **Doubly robust**: DR enjoys the property of double robustness; In contrast, IPS and EIB do not meet the property of double robustness.

- **Robust to small propensities**: Both the IPS and DR use $1/\hat{p}_{u,i}$ as the weight to recover the target distribution. In the presence of small propensities, the weights will become extremely large and cause instability. In contrast, EIB does not suffer from such a problem.

- **Boundedness**: Both the IPS and DR may lie outside the range of $L_{ideal}(\phi)$, i.e., they do not enjoy the property of boundedness. For example, if we set $e_{u,i} \in [0,1]$, then $L_{ideal}(\phi) \in [0,1]$, while $L_{IPS}(\phi)$ and $L_{DR}(\phi, \theta)$ may not be within the range. The EIB can guarantee boundedness property easily if the error imputation model is chosen appropriately.

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, "TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations." ICLR 23.

Five Desired Properties

- **Without extrapolation** (small bias): EIB usually has a large bias, which is a consequence of making implicitly extrapolation.
- Specifically, the error imputation model is trained with exposed events while using the predicted values for unexposed events.
- This relies heavily on extrapolation since the exposed events are sparse and there may exist a significant difference between the distributions of exposed events and unexposed events.
- Thus, it is hard to obtain accurate error imputation and leads to poor performance.
- In comparison, the estimation of propensity score doesn't rely on extrapolation.

- **Low variance**: It can be shown that EIB has the smallest variance among these methods.

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, “TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations.” ICLR 23.

A high-level perspective of debiasing methods in RS

Model Evaluation
(construct estimator of ideal loss)

Dynamic updating prediction model, propensity model, imputation model.

Model Learning
(algorithm and model architecture design)

PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
    - Bias-Variance Trade-Off
    - Robust to Small Propensities (Data Sparsity)
    - Robust to Pseudo-Labelings
    - Mitigating/Eliminating Unmeasured Confounding
    - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Overview

- **Bias-Variance Trade-Off**: MRDR, DR-MSE
- **Data Sparsity** (robust to small propensities): ESMM, Multi-DR, ESCM2-DR, SDR
- **Robust to Pseudo-Labelings**: MR, TDR
- **Mitigating/Eliminating Unmeasured Confounding**: BRD, BAL-IPS, BAL-DR
- **How to Set Proper Propensity**: LTD, AutoDebias, DR-V2

Bias-Variance Trade-Off

More Robust Doubly Robust (MRDR)

MRDR enhances the robustness of DR-JL by optimizing the variance of the DR estimator with the imputation model.

$$\mathcal{L}_{DR}(\phi,\theta) = \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\left[\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i}-\hat{e}_{u,i})}{\hat{p}_{u,i}}\right],$$

DR-JL

- given $\hat{\theta}$, ϕ is updated by minimizing $\mathcal{L}_{DR}(\phi,\hat{\theta})$;
- given $\hat{\phi}$, θ is updated by minimizing

$$\mathcal{L}_e^{DR-JL}(\phi,\theta) = \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(\hat{e}_{u,i}-e_{u,i})^2}{\hat{p}_{u,i}}.$$

MRDR

- given $\hat{\theta}$, ϕ is updated by minimizing $\mathcal{L}_{DR}(\phi,\hat{\theta})$;
- given $\hat{\phi}$, θ is updated by minimizing

$$\mathcal{L}_e^{MRDR}(\theta) = \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(\hat{e}_{u,i}-e_{u,i})^2}{\hat{p}_{u,i}} \cdot \boxed{\frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}}}.$$

MRDR substitutes the loss function of the imputation model.

Siyuan Guo, Lixin Zou, Yiding Liu, Wenwen Ye, Suqi Cheng, Shuaiqiang Wang, Hechang Chen, Dawei Yin, and Yi Chang, "Enhanced Doubly Robust Learning for Debiasing Post-Click Conversion Rate Estimation". SIGIR 21.

More Robust Doubly Robust (MRDR)

This substitution can help reduce the variance of $L_{DR}(\phi, \theta)$ and hence a more variance-robust estimator might be achieved.

$$\mathbb{V}_{\mathcal{O}}[\mathcal{L}_{DR}(\phi, \theta)] = \frac{1}{|\mathcal{D}|^2}\mathbb{E}_{\mathcal{O}}[\underbrace{\sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(1 - p_{u,i})(\hat{e}_{u,i} - e_{u,i})^2}{\hat{p}_{u,i}^2}}_{\mathcal{L}_e^{MRDR}(\theta)}].$$

Siyuan Guo, Lixin Zou, Yiding Liu, Wenwen Ye, Suqi Cheng, Shuaiqiang Wang, Hechang Chen, Dawei Yin, and Yi Chang, "Enhanced Doubly Robust Learning for Debiasing Post-Click Conversion Rate Estimation". SIGIR 21.

A Generalized DR Learning Framework

SIGIR 2023 TAIPEI | TAIWAN

Existing DR methods follow the same learning framework:

The underlying loss is $L(\hat{R}, R^o) + Metric\{L(\hat{R}, R^o)\}$

Prediction model: optimize the doubly robust loss $L(\hat{R}, R^o)$ (**Error Term**)

Error imputation model: optimize some property of the DR loss $Metric\{L(\hat{R}, R^o)\}$

$$\mathcal{L}_{DR}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left[\frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}} + (1 - \frac{o_{u,i}}{\hat{p}_{u,i}}) \hat{e}_{u,i} \right]$$

$$\mathcal{L}_e(\theta, \phi) = \sum_{u,i\in\mathcal{O}} \frac{(\hat{e}_{u,i} - e_{u,i})^2}{\hat{p}_{u,i}}$$

Table 1: Generalized framework of various DR methods

	Method	Metric	Goal
	DR-JL	$\sum_{(u,i)\in\mathcal{D}} (\hat{e}_{u,i} - e_{u,i})^2$	Control error of imputation.
	MRDR	$\mathbb{V}_{\mathcal{O}}[\mathcal{L}_{DR}(\hat{\mathbf{R}}, \mathbf{R}^o)]$	Control variance.
Ours	DR-BIAS	$Bias[\mathcal{L}_{DR}(\hat{\mathbf{R}}, \mathbf{R}^o)]$	Further reduce bias.
Ours	DR-MSE	$MSE[\mathcal{L}_{DR}(\hat{\mathbf{R}}, \mathbf{R}^o)]$	Bias-variance trade-off.

$$\mathcal{L}_{ideal}(\hat{\mathbf{R}}^*) \le \underbrace{\boxed{\mathcal{L}_{DR}(\hat{\mathbf{R}}^*)}}_{\text{Error Term}} + \underbrace{\boxed{\frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{|p_{u,i} - \hat{p}_{u,i}|}{\hat{p}_{u,i}} |\hat{e}_{u,i} - e^*_{u,i}|}}_{\text{Bias Term}} + \underbrace{\boxed{\sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} (\frac{\hat{e}_{u,i} - e^*_{u,i}}{\hat{p}_{u,i}})^2}}}_{\text{Variance Term}}$$

Importantly, the proposed framework provides a valuable opportunity to develop a series of new unbiased CVR estimators with different characteristics to accommodate different application scenarios.

Quanyu Dai, Haoxuan Li, Peng Wu, Zhenhua Dong, Xiao-Hua Zhou, Rui Zhang, Xiuqiang He, Rui Zhang, and Jie Sun, “A Generalized Doubly Robust Learning Framework for Debiasing Post-Click Conversion Rate Prediction”. KDD 22.

Proposed Methods

DR-BIAS: Bias Reduced Doubly Robust Estimator

$$\mathcal{L}_e^{DR-BIAS}(\theta) = \sum_{(u,i)\in\mathcal{D}} \boxed{\boxed{\boxed{\frac{o_{u,i}(\hat{e}_{u,i}-e_{u,i})^2}{\hat{p}_{u,i}}}_{\text{DR-JL}} \cdot \frac{(1-\hat{p}_{u,i})}{\hat{p}_{u,i}}}_{\text{MRDR}} \cdot \frac{(1-\hat{p}_{u,i})}{\hat{p}_{u,i}}}_{\text{DR-BIAS}}$$

$$\begin{cases} \frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}} > 1, & \text{if } \hat{p}_{u,i} < 1/2, \\ \frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}} < 1, & \text{if } \hat{p}_{u,i} > 1/2. \end{cases}$$

This new loss Increases the penalty of the clicked events with low propensity, and decreases the penalty with high propensity.

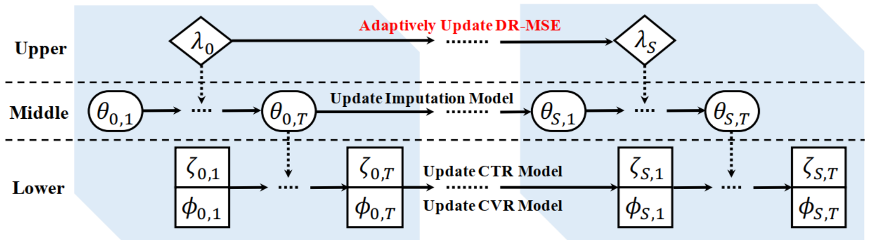
DR-MSE: Mean Square Error (MSE) Reduced Doubly Robust Estimator

$$\mathcal{L}_e^{DR-MSE}(\theta) = \lambda \underbrace{\sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(\hat{e}_{u,i}-e_{u,i})^2}{\hat{p}_{u,i}} \cdot \frac{(o_{u,i}-\hat{p}_{u,i})^2}{\hat{p}_{u,i}^2}}_{\text{Bias Term}} + (1-\lambda) \underbrace{\sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(\hat{e}_{u,i}-e_{u,i})^2}{\hat{p}_{u,i}} \cdot \frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}}}_{\text{Variance Term}}$$

- Achieve a balance between the bias term and variance to improve generalization
- Propose a tri-level optimization problem to enable adaptive bias and variance trade-off

Generalization error bound

$$\mathcal{L}_{ideal}(\hat{\mathbf{R}}^*) \leq \underbrace{\mathcal{L}_{DR}(\hat{\mathbf{R}}^*)}_{\text{Error Term}} + \underbrace{\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} \frac{|p_{u,i}-\hat{p}_{u,i}|}{\hat{p}_{u,i}} |\hat{e}_{u,i}-e^*_{u,i}|}_{\text{Bias Term}} + \underbrace{\sqrt{\frac{\log(2|\mathcal{H}|/\eta)}{2|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} (\frac{\hat{e}_{u,i}-e^*_{u,i}}{\hat{p}_{u,i}})^2}}_{\text{Variance Term}}$$



Quanyu Dai, Haoxuan Li, Peng Wu, Zhenhua Dong, Xiao-Hua Zhou, Rui Zhang, Xiuqiang He, Rui Zhang, and Jie Sun, "A Generalized Doubly Robust Learning Framework for Debiasing Post-Click Conversion Rate Prediction". KDD 22.

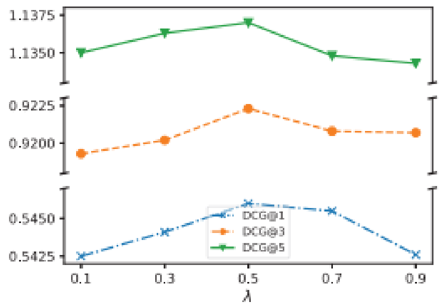
Experiments

- Performance comparison
  - Real word datasets: **Coat** and **Yahoo**
  - Industrial dataset: **Product**
  - Semi-synthetic dataset: **ML100K**
- Evaluation protocol: DCG@K and Recall@K
- Study of DR-MSE
  - Bias and variance trade-off (λ)
  - Sample ratio of unobserved data (sample ratio)

Table 3: Performance comparison based on Coat and Yahoo.

Datasets	Models	DCG@2	DCG@4	DCG@6	Recall@2	Recall@4	Recall@6
Coat	Naïve	0.7283 ± 0.0264	0.9763 ± 0.0258	1.1512 ± 0.0241	0.8474 ± 0.0310	1.3786 ± 0.0374	1.8490 ± 0.0379
	IPS	0.7102 ± 0.0220	0.9596 ± 0.0222	1.1299 ± 0.0210	0.8248 ± 0.0272	1.3596 ± 0.0360	1.8174 ± 0.0377
	DR-JL	0.7416 ± 0.0224	1.0021 ± 0.0224	1.1762 ± 0.0229	0.8645 ± 0.0264	1.4225 ± 0.0362	1.8906 ± 0.0403
	MRDR	0.7442 ± 0.0225	1.0132 ± 0.0219	1.1947 ± 0.0194	0.8736 ± 0.0273	1.4494 ± 0.0325	1.9370 ± 0.0318
	DR-BIAS	**0.7648 ± 0.0192***	**1.0353 ± 0.0169***	**1.2127 ± 0.0162***	**0.8959 ± 0.0251***	**1.4751 ± 0.0273***	**1.9517 ± 0.0324***
	DR-MSE	**0.7682 ± 0.0151***	**1.0401 ± 0.0150***	**1.2170 ± 0.0139***	**0.8997 ± 0.0194***	**1.4816 ± 0.0241***	**1.9569 ± 0.0262***
Yahoo	Naïve	0.5469 ± 0.0009	0.7466 ± 0.0008	0.8714 ± 0.0004	0.6479 ± 0.0012	1.0745 ± 0.0016	1.4098 ± 0.0013
	IPS	0.5502 ± 0.0010	0.7520 ± 0.0009	0.8751 ± 0.0009	0.6545 ± 0.0017	1.0797 ± 0.0017	**1.4168 ± 0.0019**
	DR-JL	0.5602 ± 0.0034	0.7586 ± 0.0030	0.8808 ± 0.0025	0.6615 ± 0.0042	1.0849 ± 0.0049	1.4129 ± 0.0039
	MRDR	0.5623 ± 0.0024	0.7603 ± 0.0027	0.8820 ± 0.0020	0.6646 ± 0.0033	1.0881 ± 0.0045	1.4145 ± 0.0037
	DR-BIAS	**0.5646 ± 0.0023***	**0.7624 ± 0.0021***	**0.8841 ± 0.0018***	**0.6676 ± 0.0026***	**1.0904 ± 0.0028***	**1.4169 ± 0.0020**
	DR-MSE	**0.5662 ± 0.0017***	**0.7639 ± 0.0016***	**0.8850 ± 0.0014***	**0.6670 ± 0.0026***	**1.0891 ± 0.0029**	1.4140 ± 0.0028

Note: * statistically significant results (p-value $\leq$ 0.05) using the paired-t-test compared with the best baseline.



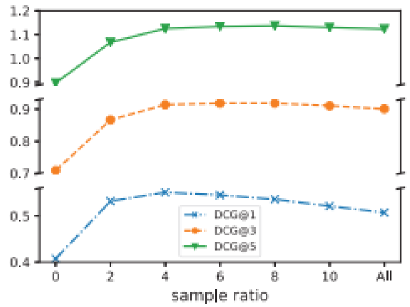


Table 4: Performance comparison based on Product.

Models	CTR AUC (%)	CVR AUC (%)	CTCVR AUC (%)
DCN	90.763	75.691	95.254
ESMM	90.704	81.647	95.505
DR-JL	90.754	81.768	95.548
Multi_IPW	90.794	81.912	95.571
Multi_DR	90.807	81.864	95.569
MRDR	90.721	81.810	95.535
DR-BIAS	**90.913**	**81.974**	**95.633**
DR-MSE	**90.825**	**82.067**	**95.654**

Table 5: Semi-synthetic datasets based on ML-100k.

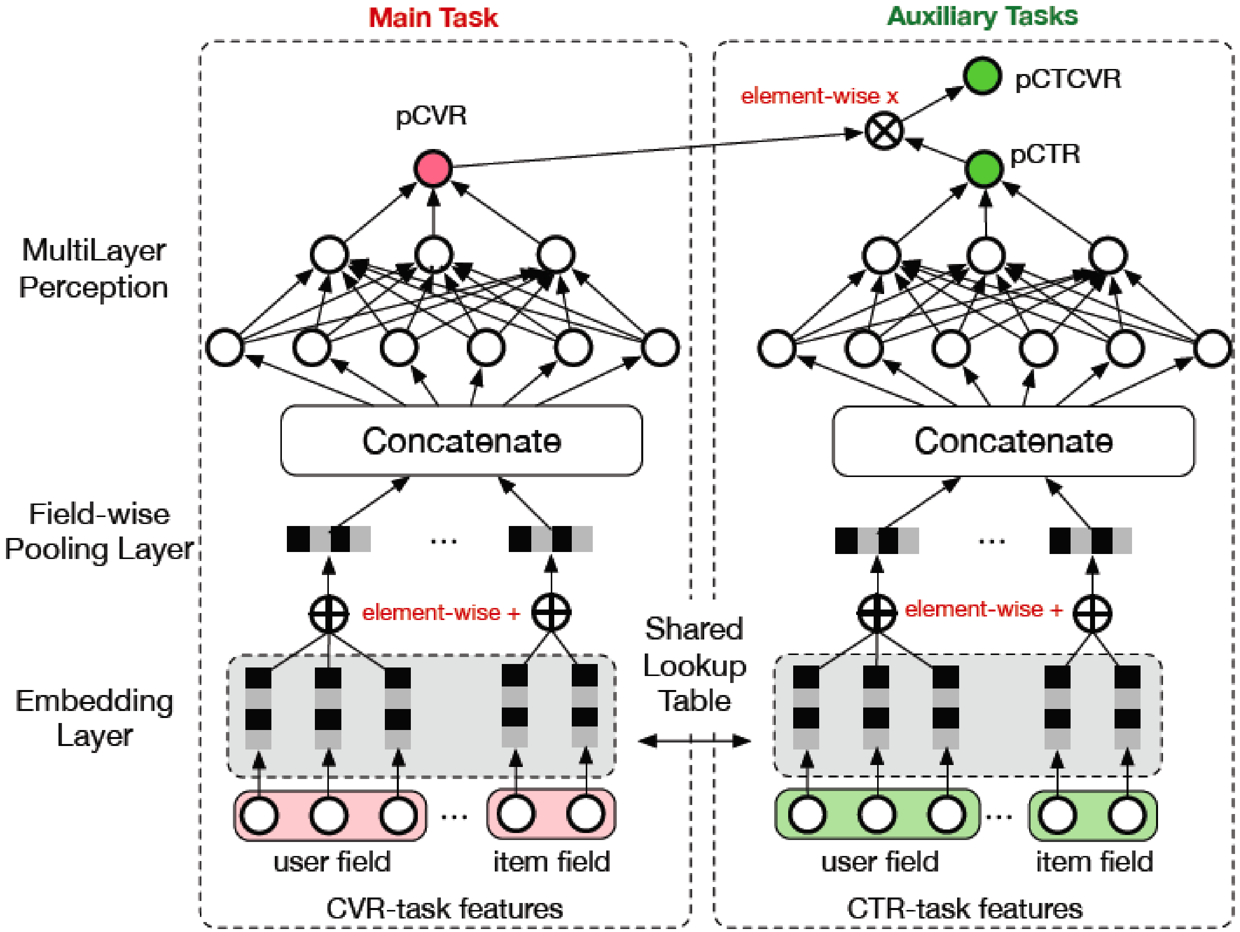
Metrics	AUC		
ρ	0.5	1	2
Naïve	0.7250 ± 0.0001	0.6731 ± 0.0001	0.5279 ± 0.0070
IPS	0.7316 ± 0.0001	0.6648 ± 0.0028	0.5263 ± 0.0055
DR-JL	0.7319 ± 0.0004	0.6673 ± 0.0035	0.5703 ± 0.0032
MRDR	0.7335 ± 0.0006	0.6765 ± 0.0021	0.5563 ± 0.0082
DR-BIAS	**0.7349 ± 0.0006***	**0.6916 ± 0.0009***	**0.6073 ±0.0054***
DR-MSE	**0.7359 ± 0.0002***	**0.6928 ± 0.0020***	**0.6084 ± 0.0168***

Quanyu Dai, Haoxuan Li, Peng Wu, Zhenhua Dong, Xiao-Hua Zhou, Rui Zhang, Xiuqiang He, Rui Zhang, and Jie Sun, “A Generalized Doubly Robust Learning Framework for Debiasing Post-Click Conversion Rate Prediction”. KDD 22.

Robust to Small Propensities (Data Sparsity)

Entire Space Multi-Task Model (ESMM)

- Intuition of Parameter Sharing
  - Training samples with all exposures for pCTR task is relatively much richer than pCVR task;
  - Thus, parameter sharing mechanism enables pCVR network to learn from un-clicked exposures and provides great help for alleviating the data sparsity trouble.



Ma, Xiao, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. "Entire space multi-task model: An effective approach for estimating post-click conversion rate." SIGIR 2018.

Multi-Task Learning: Multi-IPS

The Multi-IPS estimator is given as

$$\mathcal{L}_{Multi.IPS}(\phi, \eta, \Phi) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{c_{ui} L(r_{ui}, f(x_{u,i}; \phi, \Phi))}{\hat{p}_{u,i}(x_{u,i}; \eta, \Phi)},$$

- $\hat{p}_{u,i} = \hat{p}_{u,i}(x_{u,i}; \eta, \Phi)$ is the propensity score model, i.e., post-view click-through rate prediction model.
- $\hat{r}_{u,i} = f(x_{u,i}; \phi, \Phi)$ is the post-click conversion rate prediction model.
- Φ represents the shared embedding parameters.

Wenhao Zhang, Wentian Bao, Xiao-Yang Liu, Keping Yang, Quan Lin, Hong Wen, Ramin Ramezani, “Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning”. WWW 2020.

Multi-Task Learning: Multi-DR

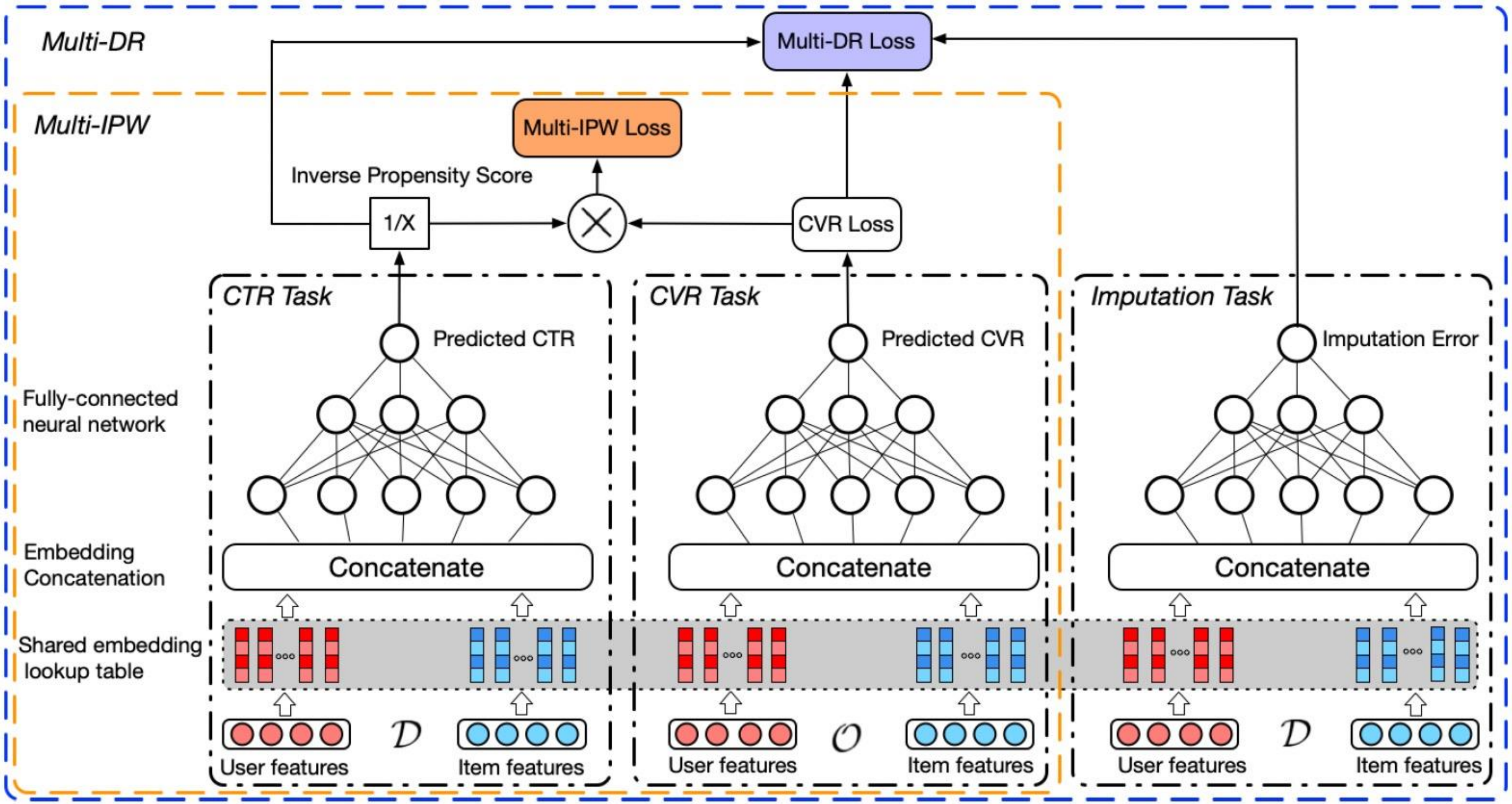
The Multi-DR estimator is given as

$$\mathcal{L}_{Multi.DR}(\phi, \eta, \Phi) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \Big\{ g_{u,i}(x_{u,i}; \theta, \Phi) + \frac{c_{ui}\left(L(r_{ui}, f(x_{u,i}; \phi, \Phi)) - g_{u,i}(x_{u,i}; \theta, \Phi)\right)}{\hat{p}_{u,i}(x_{u,i}; \eta, \Phi)} \Big\},$$

- $g_{u,i}(x_{u,i}; \theta, \Phi)$ is the error imputation model.
- $\hat{p}_{u,i} = \hat{p}_{u,i}(x_{u,i}; \eta, \Phi)$ is the propensity score model, i.e., post-view click-through rate prediction model.
- $\hat{r}_{u,i} = f(x_{u,i}; \phi, \Phi)$ is the post-click conversion rate prediction model.
- Φ represents the shared embedding parameters among CTR task, CVR task, and imputation task.

Wenhao Zhang, Wentian Bao, Xiao-Yang Liu, Keping Yang, Quan Lin, Hong Wen, Ramin Ramezani, "Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning". WWW 2020.

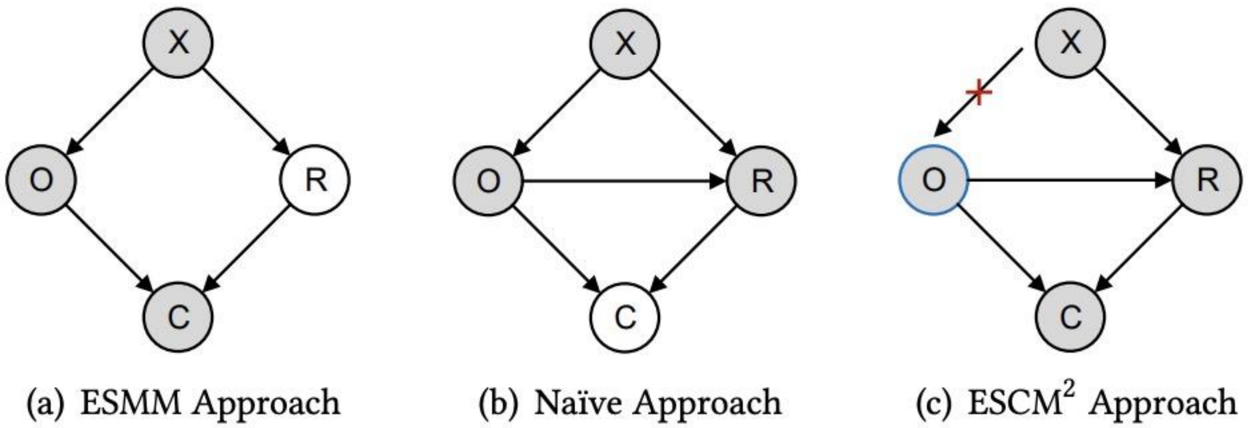
Multi-Task Learning: Multi-IPS and Multi-DR



Wenhao Zhang, Wentian Bao, Xiao-Yang Liu, Keping Yang, Quan Lin, Hong Wen, Ramin Ramezani, "Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning". WWW 2020.

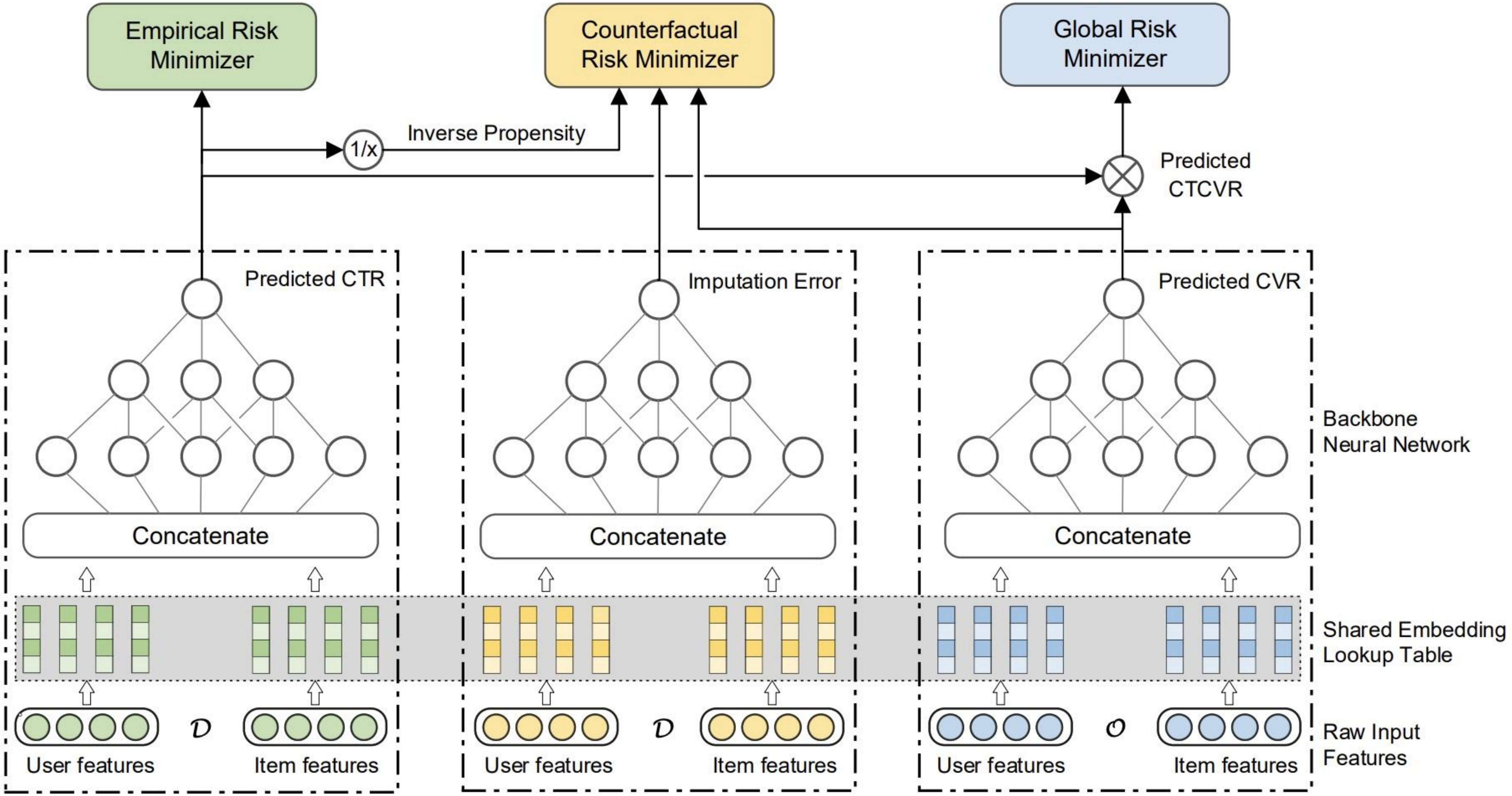
ESCM2 : Entire Space Counterfactual Multi-Task Model

- This work rigorously demonstrates the inherent bias of ESMM's CVR estimates. Mathematical proofs and experiment results are provided to support this claim.
- Show that the ESMM's CTCVR estimates are subjected to potential independence priority (PIP), also have designed experiments to back up this claim.
- Propose ESCM2, improves ESMM from a causal perspective. ESCM2 effectively eliminates Inherent Estimation Bias (IEB) and PIP in ESMM. Extensive experimental results and mathematical proofs are provided to verify the claims.



Wang, Hao, Tai-Wei Chang, Tianqiao Liu, Jianmin Huang, Zhichao Chen, Chao Yu, Ruopeng Li, and Wei Chu. "ESCM2: Entire space counterfactual multi-task model for post-click conversion rate estimation." SIGIR 2022.

ESCM2 : Entire Space Counterfactual Multi-Task Model



Wang, Hao, Tai-Wei Chang, Tianqiao Liu, Jianmin Huang, Zhichao Chen, Chao Yu, Ruopeng Li, and Wei Chu. "ESCM2: Entire space counterfactual multi-task model for post-click conversion rate estimation." SIGIR 2022.

StableDR: Stabilized Doubly Robust Learning

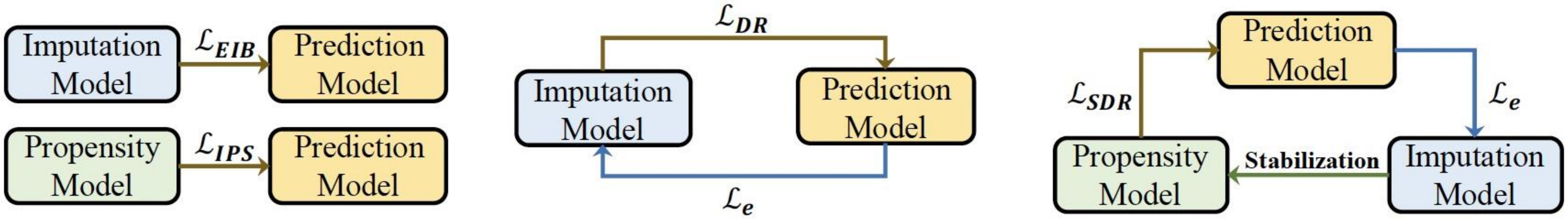
- In this paper, the authors show that DR methods are unstable and have unbounded bias, variance, and generalization bounds to extremely small propensities.
- Moreover, the fact that DR relies more on extrapolation will lead to suboptimal performance.
- To address the above limitations while retaining double robustness, we propose a stabilized doubly robust (StableDR) learning approach with a weaker reliance on extrapolation.
- Theoretical analysis shows that StableDR has bounded bias, variance, and generalization error bound simultaneously under inaccurate imputed errors and arbitrarily small propensities.
- In addition, we propose a novel learning approach for StableDR that updates the imputation, propensity, and prediction models cyclically, achieving more stable and accurate predictions.

Haoxuan Li, Chunyuan Zheng, Peng Wu, “StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random.” ICLR 23.

StableDR: Stabilized Doubly Robust Learning

	IPS	DR	SDR
Extrapolation	No (propensity model doesn't require)	Yes (due to the imputation model in DR)	Weaker than DR
Bias Robust to small $\hat{p}_{u,i}$	$\lvert\mathcal{D}\rvert^{-1} \lvert \sum_{u,i\in\mathcal{D}} (\hat{p}_{u,i} - p_{u,i}) e_{u,i}/\hat{p}_{u,i} \rvert$ No, Bias $(\mathcal{L}_{\mathrm{IPS}}) \to \infty$ when $\hat{p}_{u,i} \to 0$	$\lvert\mathcal{D}\rvert^{-1} \lvert \sum_{u,i\in\mathcal{D}} (\hat{p}_{u,i} - p_{u,i}) e_{u,i}/\hat{p}_{u,i} \rvert$ No, Bias $(\mathcal{L}_{\mathrm{DR}}) \to \infty$ when $\hat{p}_{u,i} \to 0$	See Theorem 2(a) Yes
Variance Robust to small $\hat{p}_{u,i}$	$\lvert\mathcal{D}\rvert^{-2} \sum_{u,i\in\mathcal{D}} p_{u,i}(1-p_{u,i}) e_{u,i}^2/\hat{p}_{u,i}^2$ No, Var $(\mathcal{L}_{\mathrm{IPS}}) \to \infty$ when $\hat{p}_{u,i} \to 0$	$\lvert\mathcal{D}\rvert^{-2} \sum_{u,i\in\mathcal{D}} p_{u,i}(1-p_{u,i}) (e_{u,i} - \hat{e}_{u,i})^2/\hat{p}_{u,i}^2$ No, Var $(\mathcal{L}_{\mathrm{DR}}) \to \infty$ when $\hat{p}_{u,i} \to 0$	See Theorem 2(b) Yes
Error Bound Robust to small $\hat{p}_{u,i}$	$\lvert \mathcal{L}_{\mathrm{IPS}} - \mathbb{E}_O [\mathcal{L}_{\mathrm{IPS}}] \rvert \leq \sqrt{\frac{\log\left(\frac{2}{\eta}\right)}{2\lvert\mathcal{D}\rvert^2} \sum_{u,i\in\mathcal{D}} \left(\frac{e_{u,i}}{\hat{p}_{u,i}}\right)^2}$ No, the error bound of IPS $\to \infty$ when $\hat{p}_{u,i} \to 0$	$\lvert \mathcal{L}_{\mathrm{DR}} - \mathbb{E}_O [\mathcal{L}_{\mathrm{DR}}] \rvert \leq \sqrt{\frac{\log\left(\frac{2}{\eta}\right)}{2\lvert\mathcal{D}\rvert^2} \sum_{u,i\in\mathcal{D}} \left(\frac{e_{u,i}}{\hat{p}_{u,i}}\right)^2}$ No, the error bound of DR $\to \infty$ when $\hat{p}_{u,i} \to 0$	See Theorem 4 Yes
Learning Approach	Two-phase Learning	Joint Learning	Cycle Learning

Note: $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i}$ is the error deviation.



(a) Two-phase learning using single model. (b) Doubly robust joint / double learning. (c) Proposed cycle learning with stabilization.

Haoxuan Li, Chunyuan Zheng, Peng Wu, "StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random." ICLR 23.

StableDR: Stabilized Doubly Robust Learning

- The proposed stabilized doubly robust (SDR) estimator that has a weaker dependence on extrapolation and is robust to small propensities.
- The SDR estimator consists of the following three steps.
- **Step 1 (Initialize imputed errors).** Pre-train imputation model $\hat{e}_{u,i}$, let $\hat{\mathcal{E}} = |\mathcal{D}|^{-1} \sum_{(u,i)\in\mathcal{D}} \hat{e}_{u,i}$.
- **Step 2 (Learn constrained propensities).** Learn a propensity model $\hat{p}_{u,i}$ satisfying

$$|\mathcal{D}|^{-1} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}} \left(\hat{e}_{u,i} - \hat{\mathcal{E}}\right) = 0.$$

- **Step 3 (SDR estimator).** The SDR estimator is given as $\mathcal{E}_{SDR} = \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}} / \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}}$.
- Specifically, the Step 2 is designed to enable double robustness property.

Haoxuan Li, Chunyuan Zheng, Peng Wu, "StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random." ICLR 23.

Robust to Pseudo-Labelings

Multiple Robust Learning

- Doubly robust (DR) learning has been studied, with the advantage that unbiased learning can be achieved when either a single imputation or a single propensity model is accurate.
- This paper proposes a multiple robust (MR) estimator that can take the advantage of multiple candidate imputation and propensity models to achieve unbiasedness.
- Specifically, the MR estimator is unbiased when any of the imputation or propensity models, or a linear combination of these models is accurate.
- Theoretical analysis shows that the proposed MR is an enhanced version of DR when only having a single imputation and propensity model, and has a smaller bias.
- Inspired by the generalization error bound of MR, the authors further propose a multiple robust learning approach with stabilization.

Haoxuan Li, Quanyu Dai, Yuru Li, Zhenhua Dong, Xiao-Hua Zhou, Peng Wu, "Multiple Robust Learning for Recommendation," AAAI 23.

Multiple Robust Learning

Consider J propensity models and K imputation models:

$$\mathcal{G} = \{\pi_1(x; \hat{\alpha}_1), \ldots, \pi_J(x; \hat{\alpha}_J)\},$$
$$\mathcal{M} = \{m_1(x; \hat{\beta}_1), \ldots, m_K(x; \hat{\beta}_K)\}.$$

Let $\hat{p}^j_{u,i} \triangleq \pi_j(x_{u,i}; \hat{\alpha}_j)$ and $\hat{m}^k_{u,i} \triangleq m_k(x; \hat{\beta}_k)$.

$$\boldsymbol{u}(x_{u,i}) = \left(1/\hat{p}^1_{u,i}, \cdots, 1/\hat{p}^J_{u,i}, \hat{m}^1_{u,i}, \cdots, \hat{m}^K_{u,i}\right)^T.$$

The proposed MR estimator is given as

$$\mathcal{E}_{MR} = |\mathcal{D}|^{-1} \sum_{(u,i)\in\mathcal{D}} \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}(\theta), \quad (1)$$

where $\hat{\boldsymbol{\eta}}(\theta)$ is the solution by minimizing

$$\sum_{(u,i)\in\mathcal{D}} o_{u,i}\{e_{u,i} - \boldsymbol{u}^T(x_{u,i}) \cdot \boldsymbol{\eta}\}^2. \quad (2)$$

Theorem 1 (Multiple Robustness). *MR is consistent*[1] *when either of the following conditions hold:*

(a) there exists a linear combination of the J inverse propensities accurate, i.e., $[\hat{\mathbf{P}}^{ln}]_{u,i} = 1/p_{u,i}$;

(b) there exists a linear combination of the K imputed errors accurate, i.e. $[\hat{\mathbf{E}}^{ln}]_{u,i} = e_{u,i}$,

where $\hat{\mathbf{P}}^{ln} = \sum_{j=1}^{J} w_j \hat{\mathbf{P}}^j$ and $\hat{\mathbf{E}}^{ln} = \sum_{k=1}^{K} v_k \hat{\mathbf{E}}^k$ are the linear combinations of $\hat{\mathbf{P}}^j$ and $\hat{\mathbf{E}}^k$.

In addition, the MR estimator $\mathcal{E}_{MR}$ is unbiased, if $\hat{\boldsymbol{\eta}}$ and $\mathcal{E}_{MR}$ are obtained through different samples.

Theorem 2 (Relation to DR). *Given one error imputation model and one propensity model, then*

(a) (Enhanced double robustness) $\mathcal{E}_{MR}$ has double robustness. Furthermore, when both the imputation model and propensity model are inaccurate, $\mathcal{E}_{MR}$ retains unbiasedness in condition that $e_{u,i}$ can be linearly represented by $\hat{m}_{u,i}$ and $1/\hat{p}_{u,i}$, but $\mathcal{E}_{DR}$ doesn't.

(b) (Equivalent Form) $\mathcal{E}_{MR} = \mathcal{E}_{DR}$ if the error imputation model is accurate.

Haoxuan Li, Quanyu Dai, Yuru Li, Zhenhua Dong, Xiao-Hua Zhou, Peng Wu, "Multiple Robust Learning for Recommendation," AAAI 23.

Multiple Robust Learning

Theorem 3 (Bias of MR). *Given the J propensity models and K imputation models, with $\hat{p}^j_{u,i} > 0$ for all (u,i) pairs, then the bias of MR estimator is given as*

$$\text{Bias}\left(\mathcal{E}_{MR}\right) = \frac{1}{|\mathcal{D}|}\left|\sum_{(u,i)\in\mathcal{D}} \underbrace{\left\{1 - p_{u,i}\sum_{j=1}^{J}\frac{w_j}{\hat{p}^j_{u,i}}\right\}}_{\text{linear combination of } 1/\pi_1,\ldots,1/\pi_J} \times \underbrace{\left\{e_{u,i} - \boldsymbol{u}^T(x_{u,i})\cdot\mathbb{E}_{\mathcal{O}}[\hat{\boldsymbol{\eta}}]\right\}}_{\text{linear combination of multiple models}}\right| + O(|\mathcal{D}|^{-1}),$$

where $\sum_{j=1}^{J} w_j/\hat{p}^j_{u,i}$ is the best linear approximation of $1/p_{u,i}$.

Theorem 5 (Generalization Error Bound). *For any finite hypothesis space of predictions $\mathcal{H} = \{\hat{\mathbf{Y}}_1,\ldots,\hat{\mathbf{Y}}_{|\mathcal{H}|}\}$, then under the conditions of Theorems 1 and 4, the MR estimator deviates from the true risk $\mathcal{E}_{ideal}(\hat{\mathbf{Y}}^\dagger)$ with given $\hat{\boldsymbol{\eta}}$ is bounded with probability $1-\delta$ by*

$$\mathcal{E}_{ideal}(\hat{\mathbf{Y}}^\dagger) \le \mathcal{E}_{MR}\left(\hat{\mathbf{Y}}^\dagger\right) + \sqrt{\frac{\log(2|\mathcal{H}|/\delta)}{2|\mathcal{D}|}}\max(\Gamma-1, M)\boxed{\|\hat{\boldsymbol{\eta}}\|_1}.$$

Algorithm 1: Alternating Multiple Robust Learning with Stabilization

Input: observed ratings $\mathbf{R}^o$, propensity models $\pi_1,\ldots,\pi_j$, and stabilization parameter λ

while *stopping criteria is not satisfied* **do**
- **for** $k \in \{1,\ldots,K\}$ **do**
  - **for** *number of steps for training the k-th imputation model* **do**
    - Sample a batch of user-item pairs $\{(u_{k_l}, i_{k_l})\}_{l=1}^{L}$ from $\mathcal{O}$;
    - Update β_k by descending along the gradient $\nabla_{\beta_k}\mathcal{L}_{e_k}(\theta,\beta_k)$
  - **end**
- **end**
- **for** *number of steps for training the prediction model* **do**
  - Sample a batch of user-item pairs $\mathcal{D}'$ from $\mathcal{D}$;
  - Obtain the rated samples in $\mathcal{D}'$ as $\{(u_m, i_m)\}_{m=1}^{M} = \mathcal{O}' \subseteq \mathcal{O}$;
  - $\boxed{\eta \leftarrow [\sum_{(u,i)\in\mathcal{O}'}\boldsymbol{u}(x_{u,i})\cdot\boldsymbol{u}^T(x_{u,i}) + \lambda I]^{-1}[\sum_{(u,i)\in\mathcal{O}'}\boldsymbol{u}(x_{u,i})\cdot e_{u,i}];}$
  - Sample a batch of user-item pairs $\{(u_n, i_n)\}_{n=1}^{N}$ from $\mathcal{D}\setminus\mathcal{D}'$;
  - Update θ by descending along the gradient $\nabla_\theta\mathcal{L}_{MR}(\theta;\alpha,\beta)$
- **end**

end

Haoxuan Li, Quanyu Dai, Yuru Li, Zhenhua Dong, Xiao-Hua Zhou, Peng Wu, "Multiple Robust Learning for Recommendation," AAAI 23.

Multiple Robust Learning

Table 1: Experimental results on Coat and Yahoo with MF and NCF as backbone models.

Datasets	Coat				Yahoo			
Methods	MSE	AUC	nDCG@5	nDCG@10	MSE	AUC	nDCG@5	nDCG@10
MF	0.2405	0.7028	0.6189	0.6858	0.2494	0.6806	0.6357	0.7640
+IPS	0.2251	0.7152	0.6256	0.6934	0.2223	0.6831	0.6480	0.7665
+SNIPS	0.2262	0.7082	0.6198	0.6861	0.1941	0.6834	0.6400	0.7648
+DR	0.2325	0.7121	0.6246	0.6938	0.2106	0.6849	0.6580	0.7738
+DR-JL	0.2312	0.7110	0.6209	0.6907	0.2175	0.6876	0.6458	0.7655
+MRDR-JL	0.2301	0.7157	0.6325	0.6970	0.2169	0.6841	0.6465	0.7683
+CVIB	0.2201	0.7247	0.6361	0.7030	0.2621	0.6856	0.6491	0.7718
+DIB	0.2334	0.7104	0.6303	0.6986	0.2494	0.6832	0.6348	0.7633
+MR (Ours)	**0.2106**	**0.7356**	**0.6697**	**0.7343**	**0.1920**	**0.6990**	**0.6709**	**0.7833**
NCF	0.2116	0.7661	0.6293	0.7019	0.3318	0.6771	0.6532	0.7722
+IPS	0.2002	0.7692	0.6362	0.7126	0.1706	0.6882	0.6630	0.7776
+SNIPS	**0.1920**	0.7700	0.6313	0.7070	0.1697	0.6893	0.6687	0.7810
+DR	0.2146	0.7523	0.6197	0.6908	0.1702	0.6890	0.6633	0.7779
+DR-JL	0.2071	0.7612	0.6193	0.7021	0.2396	0.6811	0.6469	0.7653
+MRDR-JL	0.2036	0.7629	0.6231	0.7011	0.2340	0.6834	0.6499	0.7681
+CVIB	0.2060	0.7661	0.6244	0.6969	0.3055	0.6748	0.6701	0.7817
+DIB	0.2030	0.7681	0.6300	0.7035	0.2849	0.7007	0.6757	0.7864
+MR (Ours)	0.1945	**0.7737**	**0.6393**	**0.7159**	**0.1676**	**0.7026**	**0.7179**	**0.8112**

* The best results are highlighted in bold.

Haoxuan Li, Quanyu Dai, Yuru Li, Zhenhua Dong, Xiao-Hua Zhou, Peng Wu, “Multiple Robust Learning for Recommendation,” AAAI 23.

Multiple Robust Learning

- Effect of Imputation Model

Table 3: Performance of the MR method on Coat under different settings of imputation models, i.e., different numbers and types.

Imputation Model	MSE	AUC	nDCG@10	Imputation Model	MSE	AUC	nDCG@10
MF	0.2295	0.7209	0.7206	MF	0.2295	0.7209	0.7206
MF, MF	0.2252	0.7243	0.7301	MF, MF	0.2252	0.7243	0.7301
MF, MF, MF	0.2232	0.7332	0.7343	NCF	0.2285	0.7230	0.7328
MF, MF, MF, MF	**0.2223**	**0.7435**	**0.7563**	NCF, NCF	**0.2093**	**0.7381**	**0.7445**
MF, MF, MF, MF, MF	0.2228	0.7421	0.7494	MF, NCF	0.2143	0.7332	0.7325

* The best results are highlighted in bold.

- Effect of Propensity Model

Table 4: Performance of the MR method under different numbers and types of propensity models on Coat dataset, where the imputation model and backbone prediction model both employ MF.

Propensity Model	MSE	AUC	nDCG@10	Propensity Model	MSE	AUC	nDCG@10
NB	0.2291	0.7219	0.7204	NB	**0.2291**	**0.7219**	0.7204
NB, NB-Uni	0.2269	0.7282	0.7322	NB, NB	0.2293	0.7216	0.7195
NB, NB-Uni, User	**0.2228**	**0.7370**	**0.7347**	NB, NB, NB	0.2293	0.7216	**0.7206**

* The best results are highlighted in bold.

Haoxuan Li, Quanyu Dai, Yuru Li, Zhenhua Dong, Xiao-Hua Zhou, Peng Wu, “Multiple Robust Learning for Recommendation,” AAAI 23.

Targeted Doubly Robust (TDR)

TABLE I: Comparison of various debiasing estimators.

	IPS	SNIPS	EIB	DR	TDR
Doubly robust	×	×	×	✓	✓
Low variance	×	×	✓	o	✓
Robust to small propensities	×	o	✓	×	✓
Without extrapolation	✓	✓	×	o	o
Boundedness	×	✓	✓	×	✓

Note: symbols ✓, o and × denote good, medium and bad, respectively.

- When the imputation model is correctly specified, EIB is the most efficient estimator, with a variance smaller than that of DR and IPS.
- DR has double robustness and has the smallest bias in practice.
- Motivation: is it possible to combine the advantages of EIB and DR in RS?

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, "TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations," ICLR 23.

Targeted Doubly Robust (TDR)

- DR and EIB are related via the "correction term". Specifically, note that

$$\mathcal{L}_{DR} = \underbrace{\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\left[o_{u,i}e_{u,i} + (1-o_{u,i})\hat{e}_{u,i}\right]}_{\mathcal{L}_{EIB}} + \underbrace{\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} o_{u,i}(e_{u,i}-\hat{e}_{u,i})\frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}}}_{\text{correction term}}$$

- If the correction term …

$$\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}} o_{u,i}(e_{u,i}-\hat{e}_{u,i})\frac{1-\hat{p}_{u,i}}{\hat{p}_{u,i}} = 0.$$

- Then the EIB would have a smaller bias and the DR would have a smaller variance.

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, "TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations," ICLR 23.

Targeted Doubly Robust (TDR)

- Assume the error imputation model can be presented as $\hat{e}_{u,i} = \varphi\{h_\phi(x_{u,i})\}$,

where h is an arbitrary function, φ is a known function, such as sigmoid, etc.

- The basic idea of TDR consists of two steps.
- **Step 1 (Initialization).** Let $\hat{e}_{u,i} = \varphi\{\hat{h}(x_{u,i})\}$ be the imputed error obtained by using any of the previous methods.
- **Step 2 (Targeting).** Update $\hat{e}_{u,i}$ by fitting an extended one-parameter model as follows $\tilde{e}_{u,i}(\eta) = \varphi\{\hat{h}(x_{u,i})\} + \eta(\frac{1}{\hat{p}_{u,i}} - 1)$,

which includes a single variable $\frac{1}{\hat{p}_{u,i}} - 1$ and the offset $\hat{h}(x_{u,i})$.

- TDR reduces both the bias and variance of DR! Model-agnostic framework!

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, "TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations," ICLR 23.

Targeted Doubly Robust (TDR)

Table 2: Mean and standard deviation of the relative error on the Naive, EIB, IPS, DR and TDR.

Dataset	Methods	ONE	THREE	FIVE	ROTATE	SKEW	CRS
ML-100K	Naive	0.0688 ± 0.0025	0.0790 ± 0.0028	0.1027 ± 0.0028	0.1378 ± 0.0011	0.0265 ± 0.0021	0.1062 ± 0.0022
	EIB	0.5442 ± 0.0016	0.5878 ± 0.0017	0.6167 ± 0.0018	0.2533 ± 0.0004	0.3584 ± 0.0007	0.1443 ± 0.0007
	IPS	0.0338 ± 0.0033	0.0390 ± 0.0037	0.0511 ± 0.0033	0.0696 ± 0.0026	0.0129 ± 0.0027	0.0526 ± 0.0026
	DR	0.0140 ± 0.0034	0.0180 ± 0.0037	0.0150 ± 0.0034	0.0401 ± 0.0016	0.0101 ± 0.0027	0.0237 ± 0.0025
	TDR	**0.0053 ± 0.0026***	**0.0035 ± 0.0025***	**0.0066 ± 0.0032***	**0.0325 ± 0.0015***	**0.0029 ± 0.0020***	**0.0193 ± 0.0025***
ML-1M	Naive	0.0682 ± 0.0007	0.0783 ± 0.0007	0.1014 ± 0.0008	0.1377 ± 0.0005	0.0256 ± 0.0007	0.1054 ± 0.0006
	EIB	0.5437 ± 0.0005	0.5872 ± 0.0005	0.6157 ± 0.0005	0.2531 ± 0.0001	0.3575 ± 0.0002	0.1442 ± 0.0001
	IPS	0.0343 ± 0.0009	0.0394 ± 0.0009	0.0508 ± 0.0009	0.0687 ± 0.0006	0.0130 ± 0.0008	0.0528 ± 0.0007
	DR	0.0130 ± 0.0009	0.0168 ± 0.0009	0.0133 ± 0.0009	0.0399 ± 0.0005	0.0090 ± 0.0008	0.0229 ± 0.0007
	TDR	**0.0054 ± 0.0009***	**0.0031 ± 0.0009***	**0.0076 ± 0.0009***	**0.0324 ± 0.0005***	**0.0031 ± 0.0008***	**0.0187 ± 0.0007***

Note: * means statistically significant results (p-value ≤ 0.001) using the paired-t-test compared with the best baseline.

Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu, "TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations," ICLR 23.

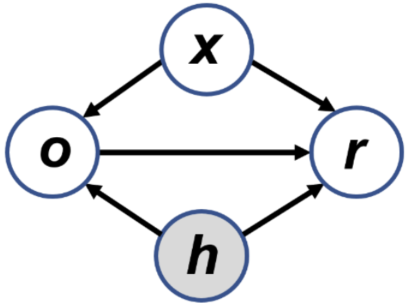
Mitigating/Eliminating Unmeasured Confounding

Benchmarked Robust Deconfounder (BRD)

- This paper reveals the risk of unmeasured confounders in recommender systems with theoretical and empirical analyses.
- The authors propose a robust deconfounding framework that mitigates unmeasured confounders with theoretical accuracy guarantee.
- Assume the nominal propensity scores are around the true ones …
- Instead of aiming to eliminate the unmeasured confounding thoroughly, the proposed RD framework calibrates the loss function with uncertainty sets by leveraging the sensitivity analysis techniques in causal inference.

Sihao Ding, Peng Wu, Fuli Feng, Yitong Wang, Xiangnan He, Yong Liao, and Yongdong Zhang, "Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis," KDD 22.

Benchmarked Robust Deconfounder (BRD)



- ***x*** : the feature of user-item pair.
- ***o*** : the exposure (treatment) status.
- ***r*** : the feedback (outcome) status.
- ***h*** : the unmeasured confounder.

Figure 2: A typical causal graph of unmeasured confounders.

THEOREM 3.1. *In the presence of unmeasured confounders h,*

(a) both the IPS and DR estimators are biased, even $\hat{p}_{u,i}$ and $\hat{e}_{u,i}$ estimate $p_{u,i}$ and $g_{u,i}$ accurately.

(b) if we define the **true** *propensity score as*

$$\boxed{\tilde{p}_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, h_{u,i}),} \tag{7}$$

and assume that $\hat{p}_{u,i}$ is an accurate estimate of $\tilde{p}_{u,i}$, then both the IPS and DR estimators are unbiased.

- **Sensitivity Analysis**

$$p_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}) = \frac{\exp(m(x_{u,i}))}{1 + \exp(m(x_{u,i}))},$$

where m is an arbitrary function. Given a bound $\Gamma \geq 1$, consider an additive model of true propensity score that

$$\tilde{p}_{u,i} = \mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, h_{u,i}) = \frac{\exp(m(x_{u,i}) + \varphi(h_{u,i}))}{1 + \exp(m(x_{u,i}) + \varphi(h_{u,i}))},$$

φ is a function and $|\varphi(h)| \leq \log(\Gamma)$, then we have

$$\frac{1}{\Gamma} \leq \frac{(1 - p_{u,i})\tilde{p}_{u,i}}{p_{u,i}(1 - \tilde{p}_{u,i})} \leq \Gamma \tag{8}$$

Eq. (8) restricts the value range of $\tilde{w}_{u,i} = 1/\tilde{p}_{u,i}$ as

$$a_{u,i} \leq \tilde{w}_{u,i} \leq b_{u,i}, \tag{9}$$

$$a_{u,i} = 1 + (1/p_{u,i} - 1)/\Gamma, \; b_{u,i} = 1 + (1/p_{u,i} - 1)\Gamma. \tag{10}$$

The hyper-parameter Γ corresponds to the strength of unmeasured confounding, and $\Gamma = 1$ means no unmeasured confounding. Let

$$\mathcal{W} = \{W \in \mathbb{R}_+^{|\mathcal{D}|} : \hat{a}_{u,i} \leq w_{u,i} \leq \hat{b}_{u,i}\}, \tag{11}$$

where $W = \{w_{u,i} : (u,i) \in \mathcal{D}\}$, $\hat{a}_{u,i}$ and $\hat{b}_{u,i}$ are the estimates of $a_{u,i}$ and $b_{u,i}$.

Sihao Ding, Peng Wu, Fuli Feng, Yitong Wang, Xiangnan He, Yong Liao, and Yongdong Zhang, "Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis," KDD 22.

Experiments

Table 2: Recommendation performances on Yahoo!R3, Coat, and Product. The best results relevant to each basic propensity-based method are highlighted with bold. RI refers to the relative improvement of RD or BRD over the corresponding baseline.

Datasets	Yahoo!R3				Coat				Product			
	UAUC	RI	NDCG@5	RI	UAUC	RI	NDCG@5	RI	UAUC	RI	NDCG@50	RI
Base model	0.6507	-	0.5449	-	0.6575	-	0.4761	-	0.6269	-	0.0914	-
DCF	0.6542	-	0.5489	-	0.6490	-	0.5016	-	0.6680	-	0.1204	-
IPS	0.6542	-	0.5525	-	0.6612	-	0.4858	-	0.6587	-	0.1131	-
RD-IPS	0.6791	3.8%	0.5808	5.1%	0.6712	1.5%	**0.5145**	5.9%	0.6680	1.4%	0.1266	12%
BRD-IPS	**0.6810**	4.1%	**0.5825**	5.4%	**0.6819**	3.1%	0.5028	3.5%	**0.6753**	2.5%	**0.1300**	15.0%
DR	0.6633	-	0.5622	-	0.6689	-	0.4949	-	0.6612	-	0.1144	-
RD-DR	0.6785	2.3%	0.5799	3.1%	**0.6803**	1.7%	**0.5092**	2.9%	0.6787	2.6%	0.1277	11.6%
BRD-DR	**0.6801**	2.5%	**0.5842**	3.9%	0.6770	1.2%	0.5080	2.8%	**0.6832**	3.3%	**0.1428**	24.8%
AutoDebias	0.7279	-	0.6421	-	0.6857	-	0.5264	-	0.6879	-	0.1365	-
RD-AutoDebias	0.7328	0.7%	0.6453	0.6%	0.6891	0.5%	0.5337	1.4%	0.6962	1.2%	**0.2183**	59.9%
BRD-AutoDebias	**0.7400**	1.7%	**0.6580**	2.6%	**0.6950**	1.4%	**0.5647**	7.3%	**0.6989**	1.6%	0.1493	9.4%

Sihao Ding, Peng Wu, Fuli Feng, Yitong Wang, Xiangnan He, Yong Liao, and Yongdong Zhang, "Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis," KDD 22.

Characters of Biased Data and Unbiased Data

- Biased data $\mathcal{D}_{\mathcal{B}}$:
  - large sample size;
  - it is inevitable to suffer from various biases.
- Unbiased data $\mathcal{D}_{\mathcal{U}}$:
  - no bias
  - it is a gold standard for evaluating the deibasing approaches.
  - small sample size, since it is costly to collect unbiased samples through uniform policy.

Only using unbiased ratings to train the rating model may cause severe overfitting due to the small sample size.

A compromised and pragmatic method is to combine two datasets: big biased observed ratings and small unbiased ratings.

Intuition of Combining Biased and Unbiased Data

- A natural question is: whether unbiased data is helpful to improve the quality of recommendations.

- Intuitively, the unbiased data provides a better way to evaluate the resulting recommendation model, and hence it may give a better-optimizing direction for training the model parameters.

- The key point is how to use the unbiased data.

- In general, unbiased data are applied to obtain a better propensity score model or error imputation (pseudo-labeling) model.

Bi-Level Optimization

- Wang et al. (2021) use the unbiased data to train the propensity score model, parameterized with η, such that the recommendation model performs well on the unbiased data.

- Formally, this goal can be formulated as a Bi-level optimization problem

$$\eta^* = \arg\min_{\eta} \mathcal{L}(\phi^*(\eta); \mathcal{D}_{\mathcal{U}})$$
$$s.t. \quad \phi^*(\eta) = \arg\min_{\phi} \mathcal{L}(\phi, \eta; \mathcal{D}_{\mathcal{B}}).$$

where

$$\mathcal{L}(\phi^*(\eta); \mathcal{D}_{\mathcal{U}}) = \sum_{(u,i)\in\mathcal{D}_{\mathcal{U}}} (r_{u,i} - f_{\phi^*(\eta)}(x_{u,i}))^2,$$

$L(\phi, \eta; D_B)$ can be chosen as the same form of IPS estimator or DR estimator.

Xiaojie Wang, Rui Zhang, Yu Sun, Jianzhong Qi, “Combating Selection Biases in Recommender Systems with a Few Unbiased Ratings”, WSDM 2021.

Mitigating Unobserved Confounding with a Few Unbiased Ratings

- Learn from uniform data:
- Uniform data provides signal on the effectiveness of debiasing.
- Meta learning mechanism:
- Base learner: optimize rec model with fixed ϕ

$$\theta^*(\phi) = \arg\min_{\theta} \sum_{(u,i)\in D_T} w_{ui}^{(1)} \delta\left(r_{ui}, \hat{r}_{ui}(\theta)\right) + \sum_{u\in U, i\in I} w_{ui}^{(2)} \delta\left(m_{ui}, \hat{r}_{ui}(\theta)\right)$$

- Meta learner: optimize debiasing parameters on uniform data

$$\phi^* = \arg\min_{\phi} \sum_{(u,i)\in D_U} \delta\left(r_{ui}, \hat{r}_{ui}(\theta^*)\right)$$

Chen, Jiawei, Hande Dong, Yang Qiu, Xiangnan He, Xin Xin, Liang Chen, Guli Lin, and Keping Yang. "AutoDebias: Learning to debias for recommendation." In SIGIR 2021.

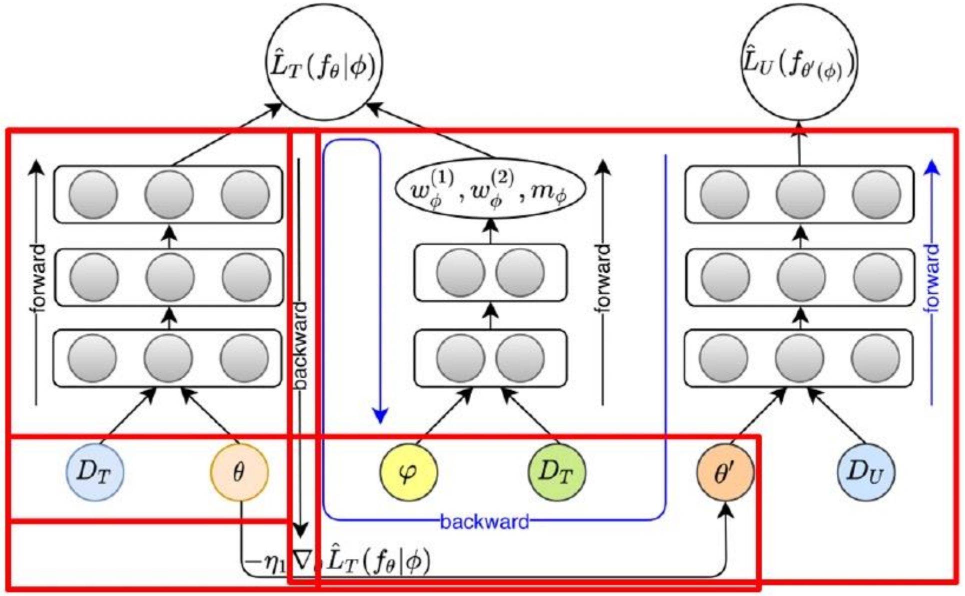
Mitigating Unobserved Confounding with a Few Unbiased Ratings

- Two challenges:
  - Overfitting: small uniform data but many debiasing parameters ϕ
    - Solution: Introduce a small meta model to generate ϕ, e.g., linear model

$$w_{ui}^{(1)} = \exp(\varphi_1^T[\mathbf{x}_u \circ \mathbf{x}_i \circ \mathbf{e}_{y_{ui}}]), \quad w_{ui}^{(2)} = \exp(\varphi_2^T[\mathbf{x}_u \circ \mathbf{x}_i \circ \mathbf{e}_{O_{ui}}]), \quad m_{ui} = \sigma(\varphi_3^T[\mathbf{e}_{y_{ui}} \circ \mathbf{e}_{O_{ui}}])$$

  - Inefficiency: obtaining optimal ϕ involves nested loops of optimization
    - Solution: Update recsys model and debiasing parameters alternately in a loop

- Step 1: Make a tentative update of θ to θ' with current ϕ
- Step 2: Test θ' on uniform data, which gives feedback to update ϕ
- Step 3: Update θ actually with updated ϕ



Balancing Unobserved Confounding with a Few Unbiased Ratings

- This paper shows the existing methods using bi-level optimization, e.g., LTD and AutoDebias, that simply uses unbiased ratings for parameter tuning of the propensity and imputation models, then the prediction models in hypothesis space are as a subset of DR.
- Though the unbiased ratings correct partial bias, in the presence of unobserved confounding or model misspecification, it is still biased due to the limited hypothesis space.

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, "Balancing Unobserved Confounding with a Few Unbiased Ratings in Debiased Recommendations," WWW 23.

Balancing Unobserved Confounding

- Motivation:

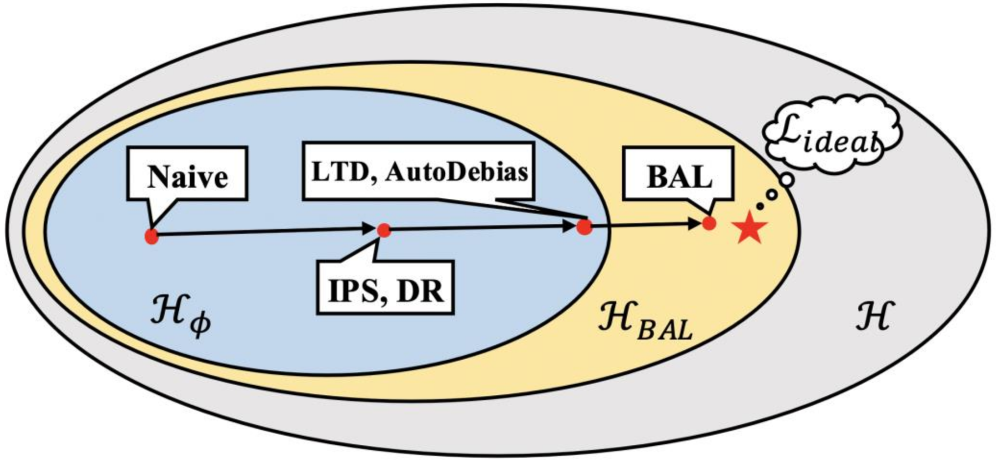


Figure 1: (i) The IPS and DR estimators learn estimates of ideal loss directly on biased ratings; (ii) LTD and AutoDebias leverage a few unbiased ratings to correct and select parameters of the propensity and imputation models, but do not enlarge the model hypothesis space, leading to biased estimates in the presence of unobserved confounding or model misspecification; (iii) The proposed model-agnostic BAL approach enlarges the hypothesis space to include the ideal loss and allows asymptotically unbiased estimation.

Proposition 2. *The IPS and DR estimators are biased, in the presence of (a) unobserved confounding or (b) model misspecification.*

Proposition 3. *(a) There exsits $w_{u,i} > 0, (u,i) \in \mathcal{B}$ such that*

$$\sum_{(u,i)\in\mathcal{B}} w_{u,i}\frac{e_{u,i}}{\hat{p}_{u,i}} = \frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}.$$

(b)There exsits $w_{u,i,1} > 0, (u,i) \in \mathcal{D}$ and $w_{u,i,2} > 0, (u,i) \in \mathcal{B}$ such that

$$\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\hat{e}_{u,i} + \sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}} = \frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}.$$

(c)There exsits $w_{u,i,1} > 0, (u,i) \in \mathcal{D}$ and $w_{u,i,2} > 0, (u,i) \in \mathcal{B}$ such that

$$\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\frac{\hat{e}_{u,i}}{\hat{p}_{u,i,1}} + \sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}}{\hat{p}_{u,i,2}} = \frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}.$$

Model-agnostic framework!

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, “Balancing Unobserved Confounding with a Few Unbiased Ratings in Debiased Recommendations,” WWW 23.

Balance Unobserved Confounding

Training Objective of Balancing Weights:

- Balanced IPS

$$\max_{\boldsymbol{w}\in\mathbb{R}^{|\mathcal{B}|}} \sum_{(u,i)\in\mathcal{B}} w_{u,i}\log(w_{u,i})$$
$$\text{s.t.}\quad w_{u,i}>0,\quad (u,i)\in\mathcal{B}$$
$$\frac{1}{|\mathcal{B}|}\sum_{(u,i)\in\mathcal{B}} w_{u,i}=\frac{1}{|\mathcal{D}|}$$
$$\sum_{(u,i)\in\mathcal{B}} w_{u,i}\frac{e_{u,i}}{\hat{p}_{u,i}}=\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i},$$

- Balanced DR and Balanced AutoDebias

$$\max_{\boldsymbol{w}_1,\boldsymbol{w}_2} \sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\log(w_{u,i,1})+\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\log(w_{u,i,2}) \tag{12}$$
$$\text{s.t.}\quad w_{u,i,1}>0,\quad (u,i)\in\mathcal{D},\qquad w_{u,i,2}>0,\quad (u,i)\in\mathcal{B} \tag{13}$$
$$\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}=1,\qquad \frac{1}{|\mathcal{B}|}\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}=\frac{1}{|\mathcal{D}|} \tag{14}$$
$$\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\hat{e}_{u,i}+\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}}=\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}, \tag{15}$$

where $\boldsymbol{w}_1=[w_{u,i,1}\mid(u,i)\in\mathcal{D}]$, $\boldsymbol{w}_2=[w_{u,i,2}\mid(u,i)\in\mathcal{B}]$, and the difference between balanced AutoDebias is that Eq. (15) comes to

$$\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\frac{\hat{e}_{u,i}}{\hat{p}_{u,i,1}}+\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}}{\hat{p}_{u,i,2}}=\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}, \tag{16}$$

- Balancing Weights Reparametrization:

$$\mathcal{L}_{W-IPS}(\xi)=-\sum_{(u,i)\in\mathcal{B}} w_{u,i}\log(w_{u,i})+\lambda\left(\sum_{(u,i)\in\mathcal{B}} w_{u,i}\frac{e_{u,i}}{\hat{p}_{u,i}}-\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}\right)^2,$$

$$\mathcal{L}_{W-DR}(\xi)=-\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\log(w_{u,i,1})-\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\log(w_{u,i,2})+\lambda\left(\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\hat{e}_{u,i}+\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}}-\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}\right)^2,$$

and

$$\mathcal{L}_{W-Auto}(\xi)=-\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\log(w_{u,i,1})-\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\log(w_{u,i,2})+\lambda\left(\sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\frac{\hat{e}_{u,i}}{\hat{p}_{u,i,1}}+\sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}}{\hat{p}_{u,i,2}}-\frac{1}{|\mathcal{U}|}\sum_{(u,i)\in\mathcal{U}} e_{u,i}\right)^2,$$

Balance Unobserved Confounding

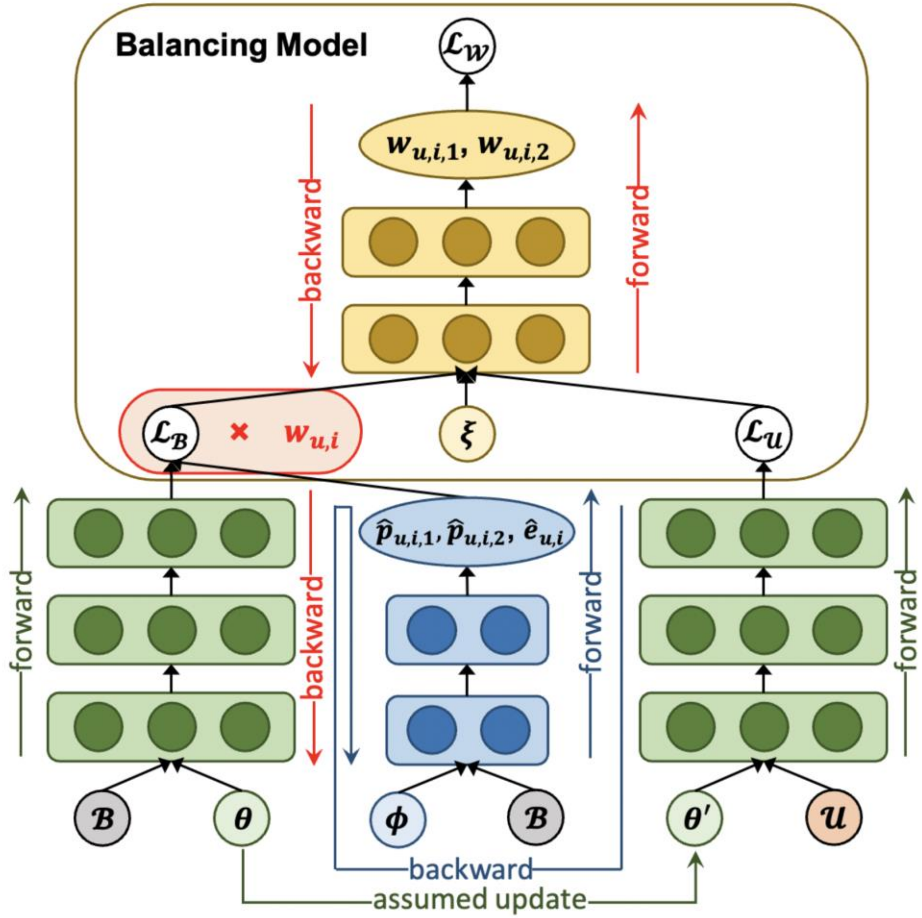
Training Objective of Prediction Model:

$$\mathcal{L}_{BAL-IPS}(\theta) = \sum_{(u,i)\in\mathcal{B}} w_{u,i}\frac{e_{u,i}}{\hat{p}_{u,i}}.$$

$$\mathcal{L}_{BAL-DR}(\theta) = \sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\hat{e}_{u,i} + \sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}},$$

and

$$\mathcal{L}_{BAL-Auto}(\theta) = \sum_{(u,i)\in\mathcal{D}} w_{u,i,1}\frac{\hat{e}_{u,i}}{\hat{p}_{u,i,1}} + \sum_{(u,i)\in\mathcal{B}} w_{u,i,2}\frac{e_{u,i}}{\hat{p}_{u,i,2}}.$$



Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, "Balancing Unobserved Confounding with a Few Unbiased Ratings in Debiased Recommendations," WWW 23.

Experiments

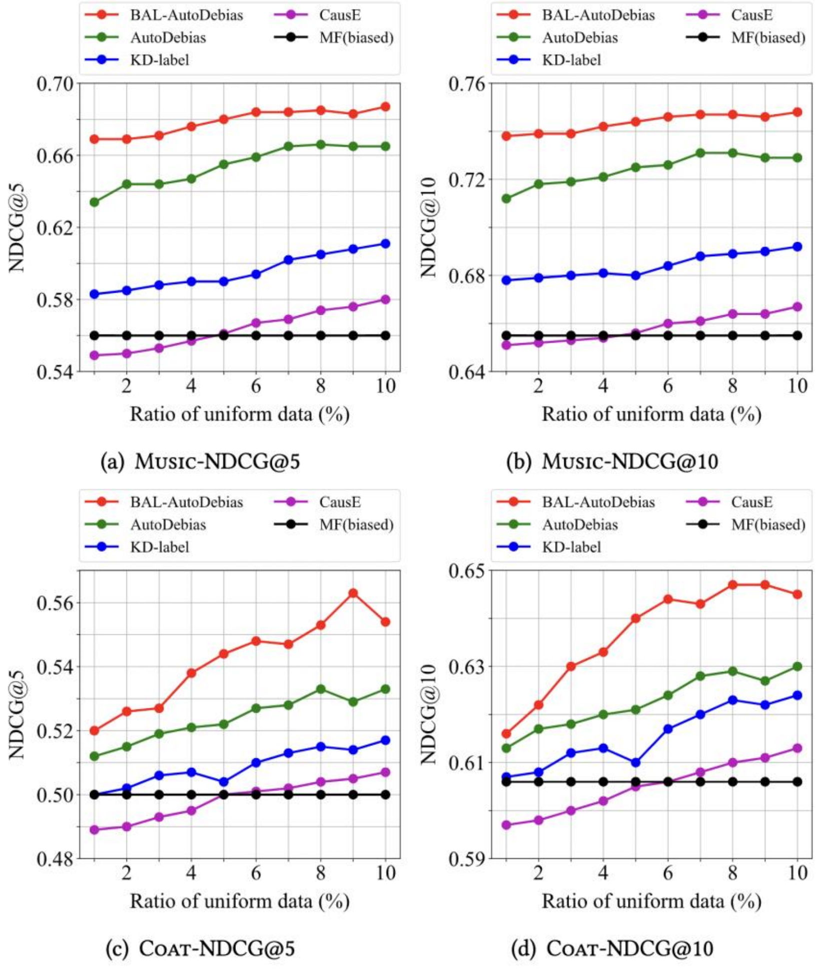
Table 2: Performance comparison in terms of AUC, NDCG@5, and NDCG@10. The best results to each base method are bolded.

Method	Music						Coat					
	AUC	RI	NDCG@5	RI	NDCG@10	RI	AUC	RI	NDCG@5	RI	NDCG@10	RI
CausE	0.731	-	0.551	-	0.656	-	0.761	-	0.500	-	0.605	-
KD-Label	0.740	-	0.580	-	0.680	-	0.750	-	0.504	-	0.610	-
MF (biased)	0.727	-	0.550	-	0.655	-	0.747	-	0.500	-	0.606	-
MF (uniform)	0.573	-	0.449	-	0.591	-	0.579	-	0.358	-	0.482	-
MF (combine)	0.730	-	0.554	-	0.659	-	0.750	-	0.503	-	0.611	-
BAL-MF	**0.739**	1.23%	**0.579**	4.51%	**0.679**	3.03%	**0.761**	1.47%	**0.511**	1.59%	**0.620**	1.47%
IPS	0.723	-	0.549	-	0.656	-	0.760	-	0.509	-	0.613	-
BAL-IPS	**0.727**	0.55%	**0.564**	2.73%	**0.668**	1.83%	**0.771**	1.45%	**0.521**	2.36%	**0.628**	2.45%
DR	0.724	-	0.550	-	0.656	-	0.765	-	0.521	-	0.620	-
BAL-DR	**0.757**	4.56%	**0.655**	19.09%	**0.729**	11.13%	**0.770**	0.65%	**0.523**	0.38%	**0.628**	1.29%
AutoDebias	0.741	-	0.645	-	0.725	-	0.766	-	0.522	-	0.621	-
BAL-AutoDebias	**0.749**	1.08%	**0.670**	3.88%	**0.744**	2.62%	**0.772**	0.78%	**0.544**	4.21%	**0.640**	3.06%

Note: RI refers to the relative improvement of BAL methods over the corresponding baseline.

Table 4: Effects of balancing models on BAL-AutoDebias.

Method		Music			Coat		
$w_{u,i,1}$	$w_{u,i,2}$	AUC	NDCG@5	NDCG@10	AUC	NDCG@5	NDCG@10
MF	MF	0.749	0.670	0.744	0.772	0.544	0.640
MF	NCF	0.745	0.667	0.742	0.769	0.539	0.635
NCF	MF	**0.762**	**0.675**	**0.748**	**0.774**	**0.548**	**0.646**
NCF	NCF	0.749	0.671	0.745	0.771	0.545	0.639



(a) Music-NDCG@5 (b) Music-NDCG@10 (c) Coat-NDCG@5 (d) Coat-NDCG@10

Figure 4: Effect of varying size of uniform data.

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, "Balancing Unobserved Confounding with a Few Unbiased Ratings in Debiased Recommendations," WWW 23.

How to Set Proper Propensity?

Motivation

- Intervene the system.
  - Position bias: randomly permutation
  - Selection bias: randomly selection

Intervene the system would harm user satisfactory.

- Inference from the observed data.
  - Training a classifier for selection or exposure.

$$P_T(u,i) = Classifier(x_u, x_i, r)$$

Approximation.

[1] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In ICML
[2] T. Joachims, A. Swaminathan, and T. Schnabel, 2017. Unbiased learning-to-rank with biased feedback. In WSDM
[3] Q. Ai, K. Bi, C. Luo, J. Guo, and W. B. Croft, 2018. Unbiased learning to rank with unbiased propensity estimation. In SIGIR.
[4] Z. Qin, S. J. Chen, D. Metzler, Y. Noh, J. Qin, and X. Wang, 2020. "Attribute-based propensity for unbiased learning in recommender systems: Algorithm and case studies. In KDD

Motivation

Despite the popularity and theoretical appeal of propensity-based approaches, a unified and clear criterion for estimating propensities has not been established yet. Many issues need to be resolved:

- How to estimate the propensity more conducive to debiasing performance?
- Which metric is more reasonable to measure the quality of the learned propensities?
- In practice, the propensities are usually trained by minimizing a cross-entropy loss. But, is it better to make the loss as small as possible when learning propensities?

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, and Peng Cui. 2023. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In ICML 23

Are NLL Proper Metrics for Propensity Model Training?

In practice, we usually train the propensity model by optimizing the cross-entropy loss (also known as the negative log-likelihood, NLL)

$$\mathcal{L}_p = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left[-o_{u,i} \log(\hat{p}_{u,i}) - (1 - o_{u,i}) \log(1 - \hat{p}_{u,i}) \right],$$

which corresponds to finding a propensity model that predicts $o_{u,i}$ as accurately as possible. However, are the learned propensities with smaller NLL sufficiently lead to a better debiasing performance?

It is obviously not. Consider an extreme case where $\hat{p}_{u,i} = 0$ for $o_{u,i} = 0$ and $\hat{p}_{u,i} = 1$ for $o_{u,i} = 1$. Although such propensities reach the smallest NLL and PLL, it would reduce $\mathcal{L}_{IPS}(\theta)$ to a Naive estimator

$$\mathcal{L}_{Naive}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} o_{u,i} e_{u,i},$$

that is, the simple averaging of losses over the observed events, which leads to biased estimates on the target population. Besides, it also reduces $\mathcal{L}_{DR}(\theta)$ to an Error Imputation-Based (EIB) estimator (Steck, 2010) that

$$\mathcal{L}_{EIB}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left[o_{u,i} e_{u,i} + (1 - o_{u,i}) \hat{e}_{u,i} \right].$$

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, and Peng Cui. 2023. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In ICML 23

Balancing-Mean-Square-Error Metric

Balancing Properties of True Propensities:

For any measurable and integrable function $\phi: \chi \rightarrow \mathbb{R}^m$, the true propensity $p_{u,i} = P(o_{u,i} = 1 | x_{u,i})$ satisfies

$$\begin{aligned}\mathbb{E}\left[\frac{o_{u,i}\phi(x_{u,i})}{p_{u,i}}\right] &= \mathbb{E}\left[\mathbb{E}\left[\frac{o_{u,i}\phi(x_{u,i})}{p_{u,i}}|x_{u,i}\right]\right] \\ &= \mathbb{E}\left[\frac{\phi(x_{u,i})}{p_{u,i}}\mathbb{E}(o_{u,i}|x_{u,i})\right] = \mathbb{E}[\phi(x_{u,i})],\end{aligned}$$

$$\mathbb{E}\left[\frac{(1-o_{u,i})\phi(x_{u,i})}{1-p_{u,i}}\right] = \mathbb{E}[\phi(x_{u,i})].$$

Balancing-Mean-Square-Error (BMSE) Metric:

$$\mathrm{BMSE}(\phi, \hat{p}) = \left|\left|\frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\left[\frac{o_{u,i}}{\hat{p}_{u,i}} - \frac{1-o_{u,i}}{1-\hat{p}_{u,i}}\right]\phi(x_{u,i})\right|\right|_F^2,$$

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, and Peng Cui. 2023. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In ICML 23

Balancing-Enhanced Estimators

The proposed balancing-enhanced IPS estimator is

$$\mathcal{L}_{IPS-V2}(\theta) = \mathcal{L}_{IPS}(\theta) + \lambda \cdot \mathrm{BMSE}(\phi, \hat{p}),$$

where $\lambda > 0$ is a scalar weight which trade-offs the balancing property and the prediction performance. Similarly, the balancing-enhanced DR estimator is

$$\mathcal{L}_{DR-V2}(\theta) = \mathcal{L}_{DR}(\theta) + \lambda \cdot \mathrm{BMSE}(\phi, \hat{p}).$$

Theorem 4.1 (Unbiasedness of IPS-V2 and DR-V2). *When learned propensities are accurate,*

(a) $\mathcal{L}_{IPS-V2}(\theta)$ is an unbiased estimator of $\mathcal{L}_{ideal}(\theta)$.

(b) $\mathcal{L}_{DR-V2}(\theta)$ is an unbiased estimator of $\mathcal{L}_{ideal}(\theta)$, whether the imputed errors are accurate or not.

Theorem 4.2 (Variance Reduction of IPS-V2 and DR-V2). *(a) Given imputed errors and learned propensities, the variance of $\mathbb{V}(\mathcal{L}_{DR-V2}(\theta) \mid \boldsymbol{o})$ reaches its minimum at*

$$\lambda_{opt} = \frac{2}{|\mathcal{D}|^2 \cdot \mathbb{V}(BMSE(\phi, \hat{p}))} \cdot \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}^2_{u,i}} \mathrm{Cov}\Bigg(e_{u,i}, \frac{1}{|\mathcal{D}|} \sum_{(s,t)\in\mathcal{D}} \left[\frac{1-o_{s,t}}{1-\hat{p}_{s,t}} - \frac{o_{s,t}}{\hat{p}_{s,t}}\right] \phi(x_{u,i})^\top \phi(x_{s,t})\Bigg),$$

where $\boldsymbol{o} = \{o_{u,i} | (u,i) \in \mathcal{D}\}$ is all the treatment indicators.

(b) $\mathcal{L}_{DR-V2}(\theta)$ has a smaller variance than $\mathcal{L}_{DR}(\theta)$,

$$\begin{aligned}\mathbb{V}(\mathcal{L}_{DR-V2}(\theta) \mid \boldsymbol{o})\Big|_{\lambda=\lambda_{opt}} &= (1-\rho^2_{L,B})\, \mathbb{V}(\mathcal{L}_{DR}(\theta) \mid \boldsymbol{o}) \\ &\le \mathbb{V}(\mathcal{L}_{DR}(\theta) \mid \boldsymbol{o}),\end{aligned}$$

where $\rho_{L,B} = \mathrm{Corr}\left(\mathcal{L}_{DR}(\theta), BMSE(\phi, \hat{p})\right)$, and similar results hold for $\mathcal{L}_{IPS-V2}(\theta)$.

Are Previous Regularizers Unbiased?

Several works have proposed estimators similar in form to the IPS-V2 and DR-V2, but with different regularization constraints (Swaminathan & Joachims, 2015b; Wang et al., 2021; Guo et al., 2021; Dai et al., 2022). For example, by using the bi-level optimization, Wang et al. (2021) adopts the sample variance (SV) regularization constraints[2]

$$\mathcal{L}_{IPS-SV}(\theta) = \mathcal{L}_{IPS}(\theta) + \lambda \cdot \mathcal{L}_{SV},$$
$$\mathcal{L}_{DR-SV}(\theta) = \mathcal{L}_{DR}(\theta) + \lambda \cdot \mathcal{L}_{SV},$$

where $\mathcal{L}_{SV} = \frac{1}{|\mathcal{D}|-1} \sum_{(u,i)\in\mathcal{D}} \left(\hat{p}_{u,i} - \frac{1}{|\mathcal{D}|} \sum_{(s,t)\in\mathcal{D}} \hat{p}_{s,t}\right)^2$

There are other alternative regularizers, such as mean inverse square (MIS) (Wang et al., 2021)

$$\mathcal{L}_{MIS} = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{1}{\hat{p}_{u,i}^2},$$

Proposition 4.3 (Bias of Previous Regularizers). *Regardless of whether the imputed errors or the learned propensities are accurate, the sample variance regularization is biased*

$$\mathbb{E}\left[\mathcal{L}_{DR-SV}(\theta)\right] = \mathbb{E}[\mathcal{L}_{DR}(\theta)] + \lambda \cdot \mathbb{E}[\mathcal{L}_{SV}] \neq \mathcal{L}_{ideal}(\theta),$$

and same for $\mathcal{L}_{IPS-SV}(\theta)$, as well as other regularizers.

Haoxuan Li, Yanghao Xiao, Chunyuan Zheng, Peng Wu, and Peng Cui. 2023. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In ICML 23

PO Framework for Recommendation

- General PO Framework
- Biases in RS and Formalization
- Debiasing Strategies: Overview
- Limitations of Basic Methods
- Enhanced Debiasing Methods
  - Bias-Variance Trade-Off
  - Robust to Small Propensities (Data Sparsity)
  - Robust to Pseudo-Labelings
  - Mitigating/Eliminating Unmeasured Confounding
  - How to Set Proper Propensity?
- Counterfactual Learning under PO Framework

Counterfactual Learning under PO Framework

Counterfactual Learning under PO Framework

- The challenge: How to identify the joint distribution of $Y_{u,i}(0)$ and $Y_{u,i}(1)$?
- Difference compared with intervention …
- Intervention problem performs the inference on subgroup level.
- Counterfactual problem performs the inference on individual level.
- Intervention problem requires the identification of $P(Y_{u,i}(0))$ and $P(Y_{u,i}(1))$.
- Counterfactual problem requires the identification of $P(Y_{u,i}(0), Y_{u,i}(1))$.

Haoxuan Li, Chunyuan Zheng, Yixiao Cao, Zhi Geng, Yue Liu, Peng Wu, "Trustworthy Policy Learning under the Counterfactual No-Harm Criterion", ICML 23.

Background

- Effective personalized incentives can improve user experience and increase platform revenue, resulting in a win-win situation between users and e-commerce companies.
- Previous studies have used uplift modeling methods to estimate the conditional average treatment effects of users' incentives, and then placed the incentives by maximizing the sum of estimated treatment effects under a limited budget.

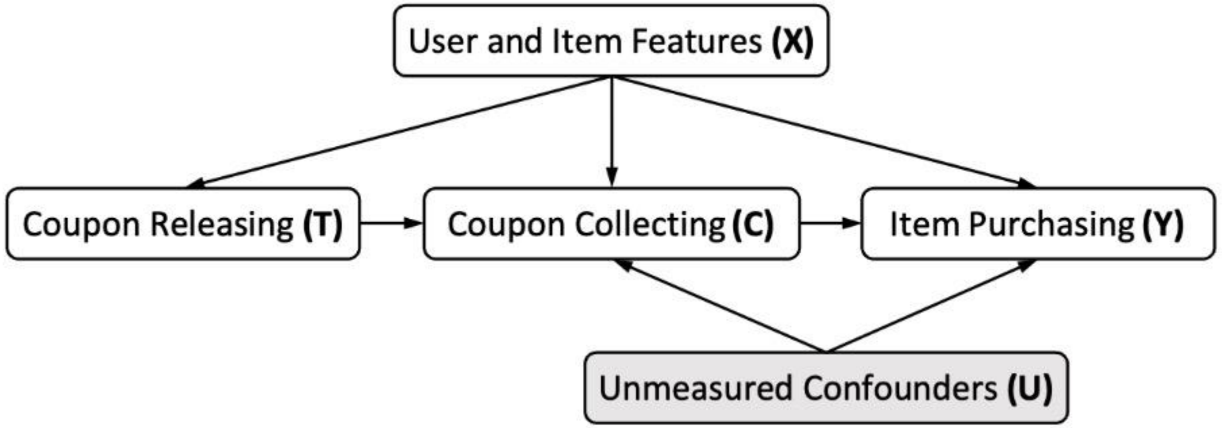


Figure 1: The causal diagram of *Coupon Releasing→Coupon Collecting→Item Purchasing* in e-commerce.

Table 1: The user-item pairs are divided into five strata from a counterfactual perspective, i.e., $(C(0), C(1), Y(0), Y(1))$, named "never buyer", "never taker", "coupon taker", "coupon buyer", and "always buyer", respectively.

Strata	Description	C(0)	C(1)	Y(0)	Y(1)	Reward
Y_{0000}	Never Buyer	0	0	0	0	0
Y_{0011}	Never Taker	0	0	1	1	0
Y_{0100}	Coupon Taker	0	1	0	0	0 or $-c(x)^*$
Y_{0101}	Coupon Buyer	0	1	0	1	$s(x)$
Y_{0111}	Always Buyer	0	1	1	1	$-c(x)$

Haoxuan Li, Chunyuan Zheng, Peng Wu, Kun Kuang, Yue Liu, Peng Cui (2023), Who should be Given Incentives? Counterfactual Optimal Treatment Regimes Learning for Recommendation. KDD 23.

Motivation

- However, some users will always buy whether incentives are given or not, and they will actively collect and use incentives if provided, named "Always Buyers".
- Identifying and predicting these "Always Buyers" and reducing incentive delivery to them can lead to a more rational incentive allocation.

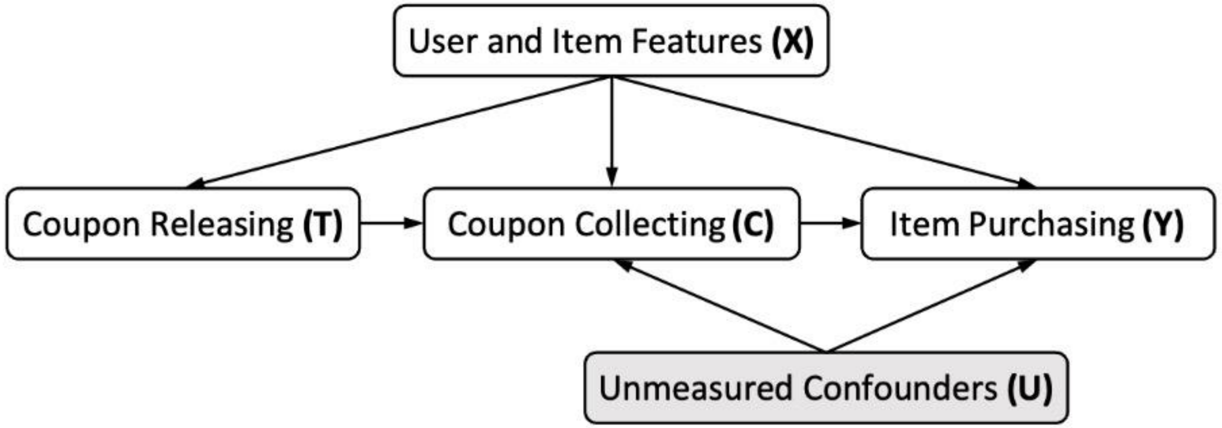


Figure 1: The causal diagram of *Coupon Releasing→Coupon Collecting→Item Purchasing* in e-commerce.

Table 1: The user-item pairs are divided into five strata from a counterfactual perspective, i.e., $(C(0), C(1), Y(0), Y(1))$, named "never buyer", "never taker", "coupon taker", "coupon buyer", and "always buyer", respectively.

Strata	Description	C(0)	C(1)	Y(0)	Y(1)	Reward
Y_{0000}	Never Buyer	0	0	0	0	0
Y_{0011}	Never Taker	0	0	1	1	0
Y_{0100}	Coupon Taker	0	1	0	0	0 or $-c(x)^*$
Y_{0101}	Coupon Buyer	0	1	0	1	$s(x)$
Y_{0111}	Always Buyer	0	1	1	1	$-c(x)$

Haoxuan Li, Chunyuan Zheng, Peng Wu, Kun Kuang, Yue Liu, Peng Cui (2023), Who should be Given Incentives? Counterfactual Optimal Treatment Regimes Learning for Recommendation. KDD 23.

Counterfactual Identification and Estimation

- First divide users into five strata from an individual counterfactual perspective, and reveal the failure of previous uplift modeling methods to identify and predict the "Always Buyers".
- Then, this paper propose principled counterfactual identification and estimation methods and prove their unbiasedness.

$$\mathbb{P}(Y_{0000} \mid X) = \mathbb{P}(C = 0, Y = 0 \mid T = 1, X),$$
$$\mathbb{P}(Y_{0011} \mid X) = \mathbb{P}(C = 0, Y = 1 \mid T = 1, X),$$
$$\mathbb{P}(Y_{0100} \mid X) = \mathbb{P}(C = 1, Y = 0 \mid T = 1, X),$$
$$\mathbb{P}(Y_{0101} \mid X) = \mathbb{P}(Y = 1 \mid T = 1, X) - \mathbb{P}(Y = 1 \mid T = 0, X),$$
$$\mathbb{P}(Y_{0111} \mid X) = \mathbb{P}(Y = 1 \mid T = 0, X) - \mathbb{P}(C = 0, Y = 1 \mid T = 1, X).$$

Haoxuan Li, Chunyuan Zheng, Peng Wu, Kun Kuang, Yue Liu, Peng Cui (2023), Who should be Given Incentives? Counterfactual Optimal Treatment Regimes Learning for Recommendation. KDD 23.

Counterfactual Entire-Space Multi-Task Learning Approach

- This paper further propose a counterfactual entire-space multi-task learning approach to accurately perform personalized incentive policy learning with a limited budget.
- Also theoretically derive a lower bound on the reward of the learned policy.

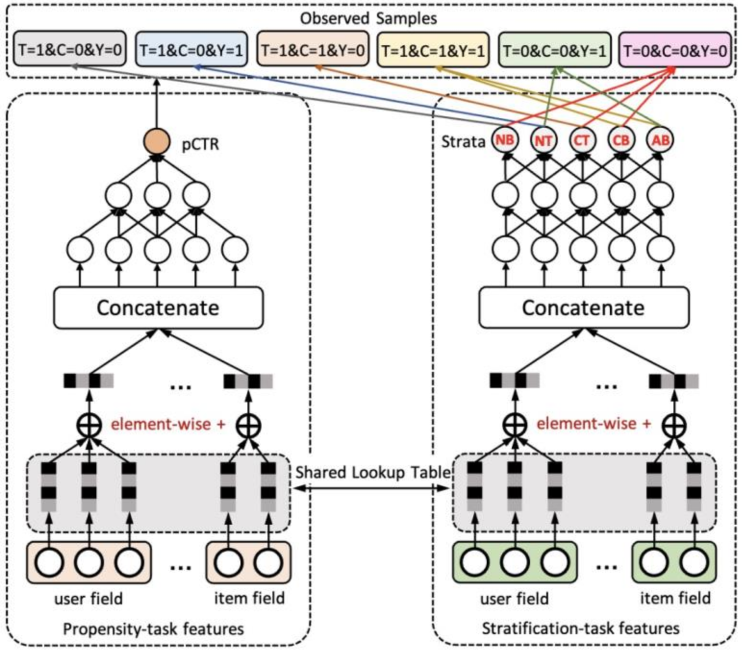


Figure 2: Proposed counterfactual entire-space multi-task learning architecture, which contains (i) a propensity model and (ii) an individual counterfactual strata prediction model.

$$\begin{aligned}
&\mathcal{L}_s(f_{0000}, f_{0011}, f_{0100}, f_{0101}, f_{0111}; g)\\
&= L\left(f_{0000}(X)g(X), T = 1\&C = 0\&Y = 0\right)\\
&+ L\left(f_{0011}(X)g(X), T = 1\&C = 0\&Y = 1\right)\\
&+ L\left(f_{0100}(X)g(X), T = 1\&C = 1\&Y = 0\right)\\
&+ L\left((f_{0101}(X) + f_{0111}(X))g(X), T = 1\&C = 1\&Y = 1\right)\\
&+ L\left((f_{0011}(X) + f_{0111}(X))(1 - g(X)), T = 0\&C = 0\&Y = 1\right)\\
&+ L\left((f_{0000}(X) + f_{0100}(X) + f_{0101}(X))(1 - g(X)), T = 0\&C = 0\&Y = 0\right)
\end{aligned}$$

$$\mathcal{L}_p(g) = L(g(X), T = 1).$$

Haoxuan Li, Chunyuan Zheng, Peng Wu, Kun Kuang, Yue Liu, Peng Cui (2023), Who should be Given Incentives? Counterfactual Optimal Treatment Regimes Learning for Recommendation. KDD 23.

Experiments

Table 4: Performance comparison of naive, uplift modeling, and proposed counterfactual learning methods, with coupon as incentive and cash as incentive on Yelp, ML-1M, and KuaiRec. We bold the best results within OR, IPS, and DR methods.

Coupon	Yelp					ML-1M					KuaiRec				
Methods	Positive	Negative	Neutral	Reward	RI	Positive	Negative	Neutral	Reward	RI	Positive	Negative	Neutral	Reward	RI
Naive	35,829	31,919	90,220	17,549	-	50,903	34,618	130,524	37,055	-	3,332	57,461	141,235	-19,652	-
OR	58,593	27,389	71,986	47,637	-	76,906	45,425	93,714	58,736	-	67,052	24,988	109,988	57,056	-
CF-OR	**58,635**	**22,557**	76,776	**49,612**	**4.14%**	**78,674**	**40,673**	96,698	**62,404**	**6.24%**	**70,366**	**23,016**	108,646	**61,159**	**7.19%**
IPS	**56,549**	26,282	75,137	46,036	-	80,035	42,770	93,240	62,927	-	82,398	17,775	101,855	75,288	-
CF-IPS	56,470	**22,933**	78,565	**47,296**	**2.73%**	**80,782**	**36,054**	99,209	**66,360**	**5.45%**	**83,694**	**16,857**	101,477	**76,951**	**2.20%**
DR	58,534	27,232	72,202	47,641	-	78,830	44,789	92,426	60,914	-	76,529	19,219	106,280	68,841	-
CF-DR	**58,757**	**22,387**	76,824	**49,802**	**4.53%**	**80,621**	**39,002**	96,422	**65,020**	**6.74%**	**78,506**	**17,346**	106,176	**71,567**	**3.95%**
CF-MTL	**67,686**	**13,397**	76,885	**62,327**	**30.82%**	**85,653**	**30,069**	100,323	**73,625**	**17.00%**	**90,538**	**11,751**	99,739	**85,837**	**14.01%**

Cash	Yelp					ML-1M					KuaiRec				
Methods	Positive	Negative	Neutral	Reward	RI	Positive	Negative	Neutral	Reward	RI	Positive	Negative	Neutral	Reward	RI
Naive	35,829	68,173	53,966	8,559	-	50,903	90,130	75,012	14,851	-	3,332	108,762	89,934	-40,172	-
OR	**58,593**	51,611	47,764	37,948	-	**76,906**	106,324	32,815	34,376	-	67,052	45,011	89,965	49,047	-
CF-OR	56,797	**40,950**	60,221	**40,417**	**6.50%**	76,747	**90,196**	49,102	**40,668**	**18.30%**	**67,171**	**44,917**	89,940	**49,204**	**0.32%**
IPS	**56,549**	49,931	51,488	36,576	-	**80,035**	93,198	42,812	42,755	-	82,398	29,816	89,814	70,471	-
CF-IPS	57,050	**39,374**	61,544	**41,300**	**12.91%**	78,636	**71,076**	66,333	**50,205**	**17.42%**	**82,451**	**29,723**	89,854	**70,561**	**0.12%**
DR	**58,534**	51,162	48,272	38,069	-	**78,830**	100,835	36,380	38,496	-	76,529	35,650	89,849	62,269	-
CF-DR	56,963	**39,120**	61,885	**41,315**	**8.52%**	78,424	**79,109**	57,512	**46,780**	**21.51%**	**76,626**	**35,499**	89,903	**62,426**	**0.25%**
CF-MTL	**67,686**	**25,549**	64,733	**57,466**	**50.95%**	**85,548**	**52,218**	78,279	**64,660**	**51.23%**	**90,187**	**21,608**	89,813	**81,963**	**16.30%**

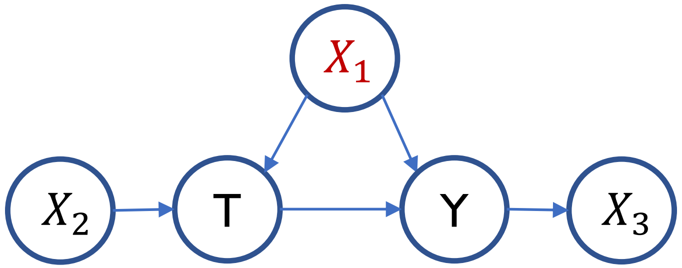
Haoxuan Li, Chunyuan Zheng, Peng Wu, Kun Kuang, Yue Liu, Peng Cui (2023), Who should be Given Incentives? Counterfactual Optimal Treatment Regimes Learning for Recommendation. KDD 23.

Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

Comparison between PO and SCM for Recommendation

- Connections
  - logically equivalent: most theorem and assumptions can be equally translated.
- SCM
  - Intuitive: use causal graph to explicitly describe causal relationships.
  - Need more knowledge and assumptions on the causal graph.
- PO
  - Easy to capture some assumptions that can not be naturally represented by DAGs, such as the identification of the Local Average Treatment Effect (LATE).



An intuitive example:

- To estimate the **causal effect of T on Y**, SCM might first assume the relationships between X_1, X_2, X_3, T, and Y, and then SCM can control X_1.
- PO might directly control X_1, X_2, and X_3 without knowing the fine-grained causal relationships.

Outline

- Part 1 (90 min, 9:00—10:30)
  - Introduction (Wenjie Wang, 15 min)
  - Structural causal models for recommendation (Yang Zhang and Wenjie Wang, 60~70 min)
  - Q&A (5 min)
  - Coffee break (30 min)
- Part 2 (90 min, 11:00-12:30)
  - Potential outcome framework for recommendation (Haoxuan Li and Peng Wu, 60~70 min)
  - Comparison (Fuli Feng, 2 min)
  - Conclusion, open problems, and future directions (Fuli Feng, 20 min)
  - Q&A (5 min)

Summary of Causal Recommendation

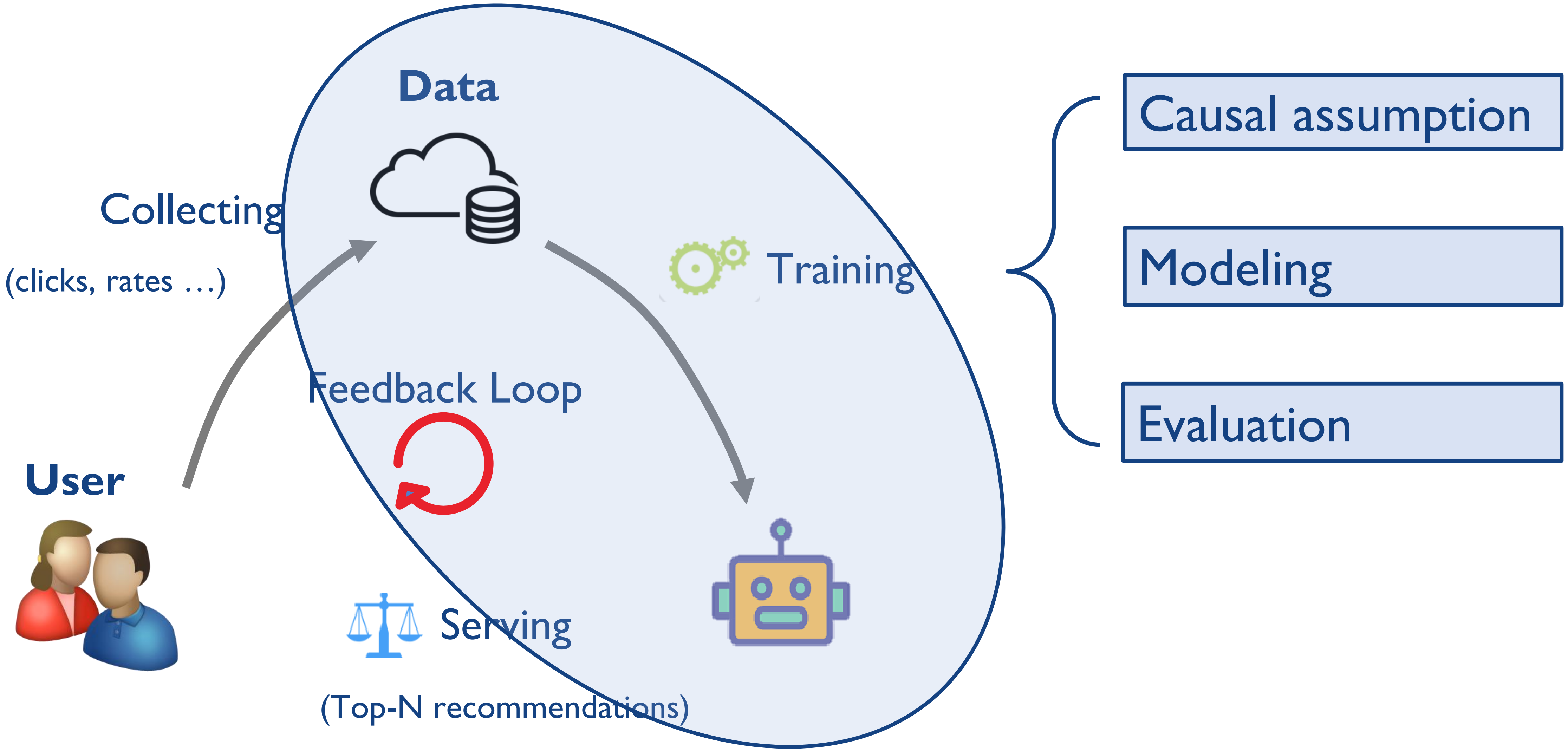
- Causal frameworks → Better recommender systems
  - Debiasing
  - Fairness
  - OOD Generalization
  - … (*Many other researches, we apologize for not covering all! Kindly let us know about your work and suggestions: wenjiewang96@gmail.com*)

- Try a causal perspective to solve your recommendation problem

- Two frameworks: PO and SCM-based methods
  - Causal graph is the key of the SCM-based methods.
  - SCM based methods may need more causal assumptions.
  - Propensity scores are usually used in PO-based methods.

- How to choose between PO and SCM? Practical requirements

Open Problems and Future Directions

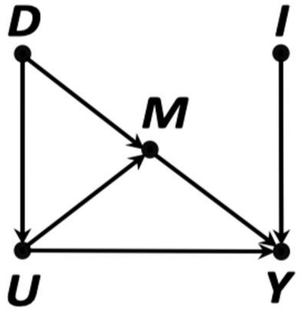


Open Problems and Future Directions

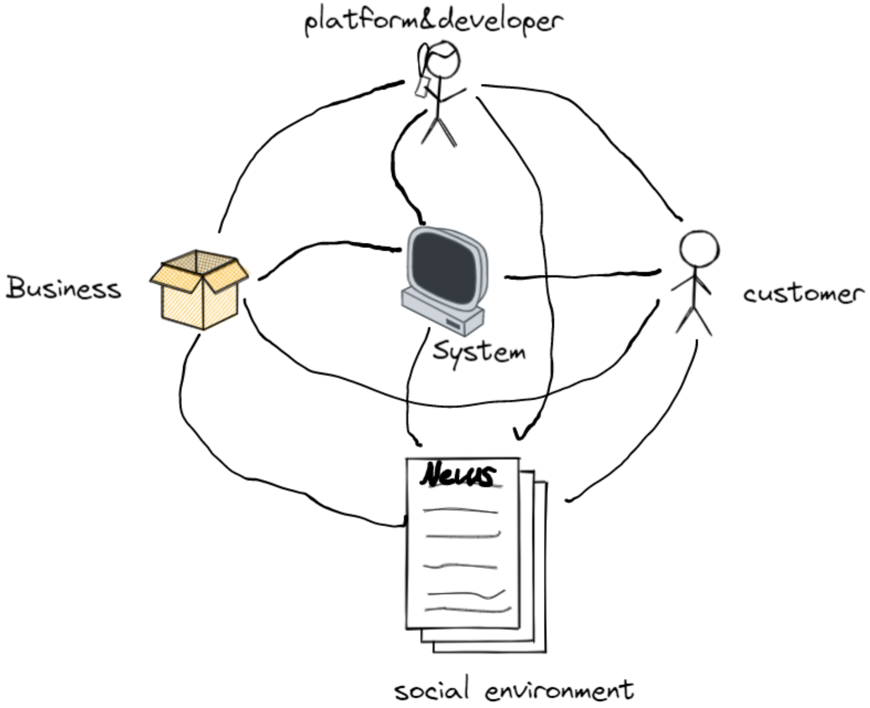
- PO & SCM requires causal assumptions
  - Existing PO-based methods need to choose covariates to satisfy the exchangeability assumption.
  - Existing SCM-based methods need to manually draw the casual graph.

$P(Y^a \perp A|L)$

POM assumption



SCM assumption

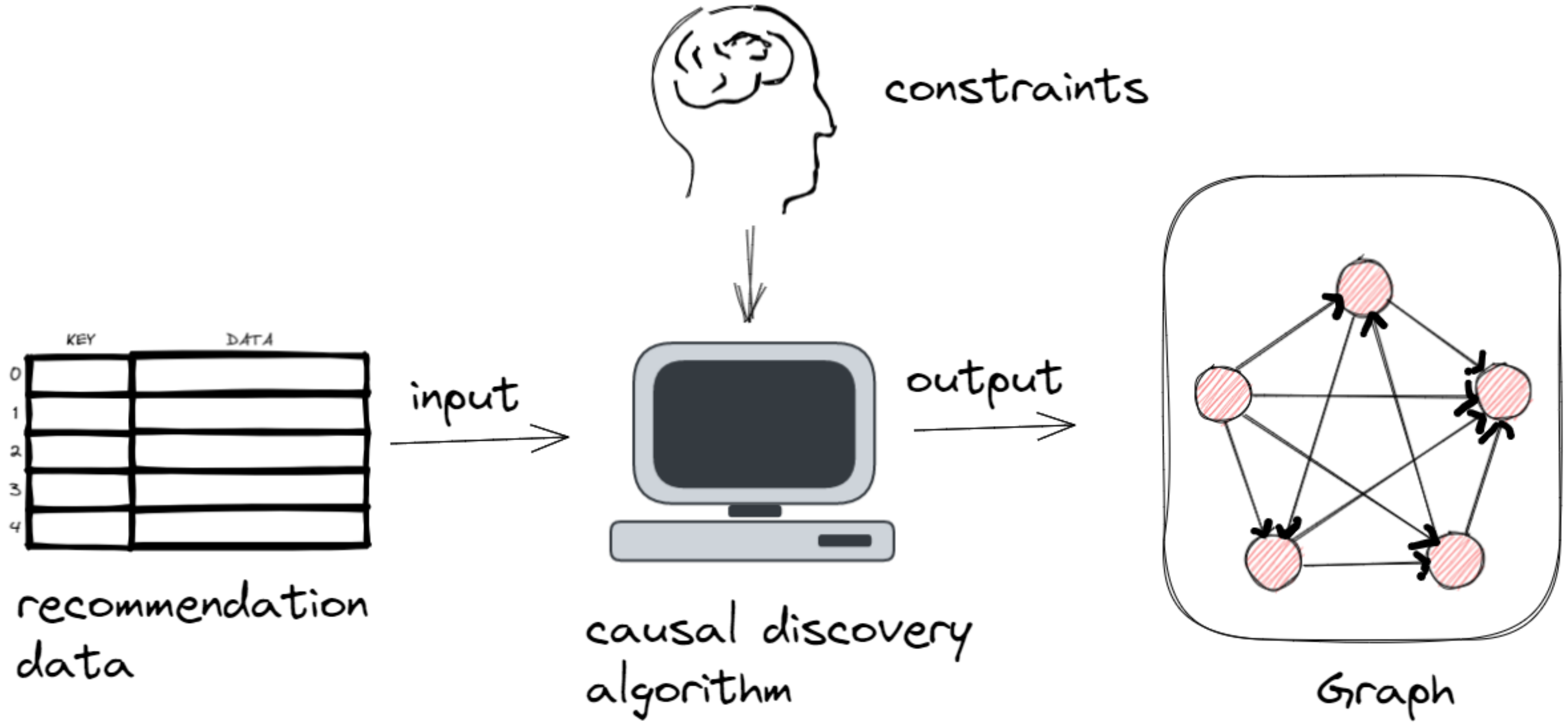


How to obtain proper causal assumptions?

- Recommender system is a complex environment.
- Prior knowledge are insufficient.

Open Problems and Future Directions

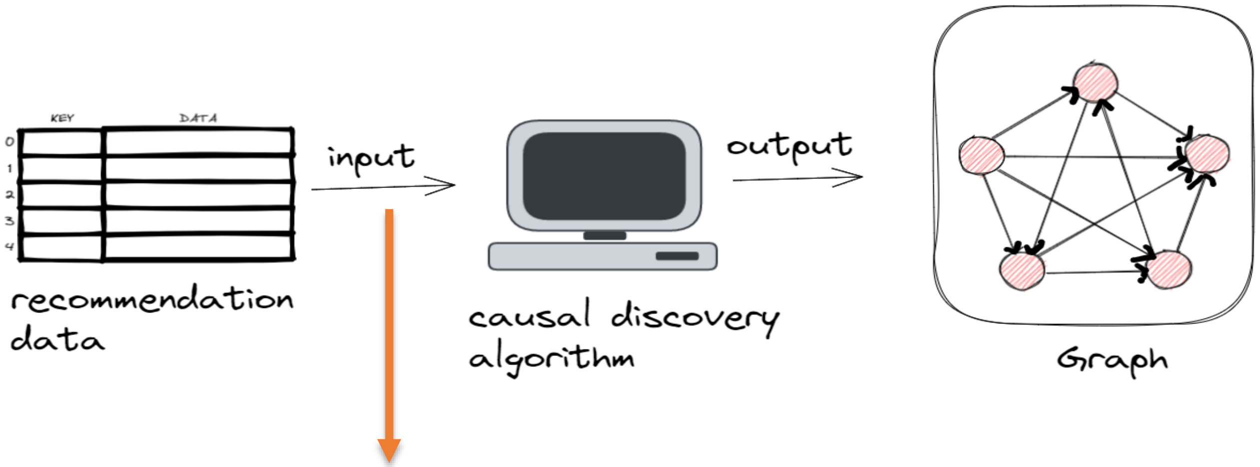
- Future direction: causal discovery in recommendation



Automatic discovery of cause graphs with causal discovery algorithms

Open Problems and Future Directions

- Future direction: causal discovery in recommendation
- Challenges for applying casual discovery algorithms in recommendation

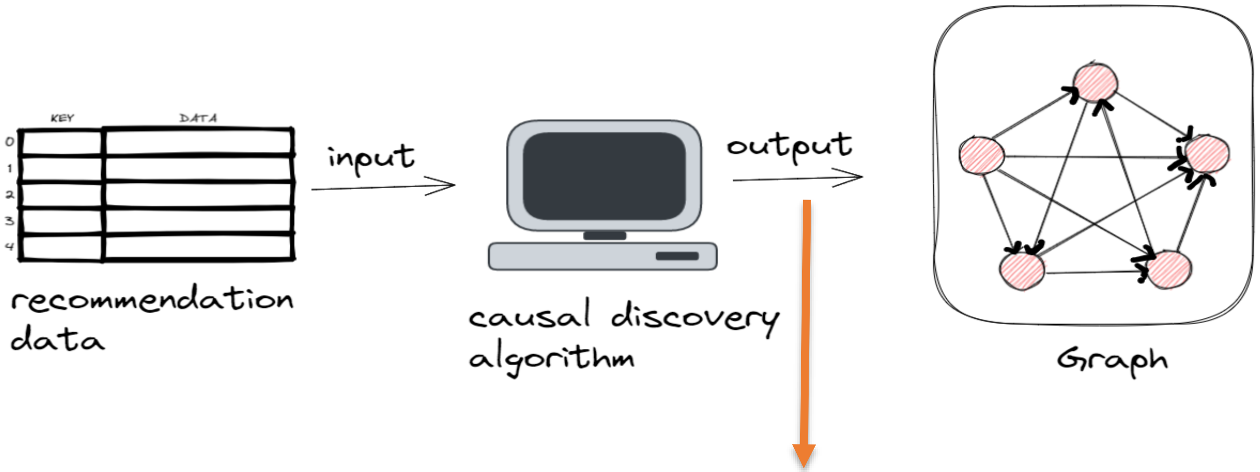


- Normal causal discovery algorithm only deals with few variables
- Challenge 1:

High-dimensional and hidden variables.

Open Problems and Future Directions

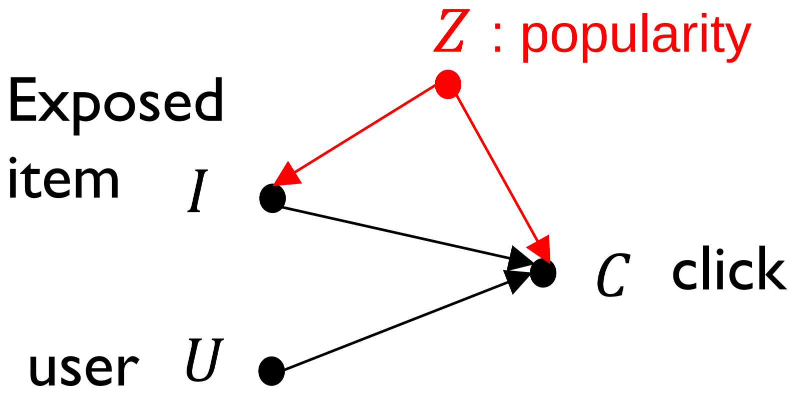
- Future direction: causal discovery in recommendation
- Challenges for applying casual discovery algorithms in recommendation



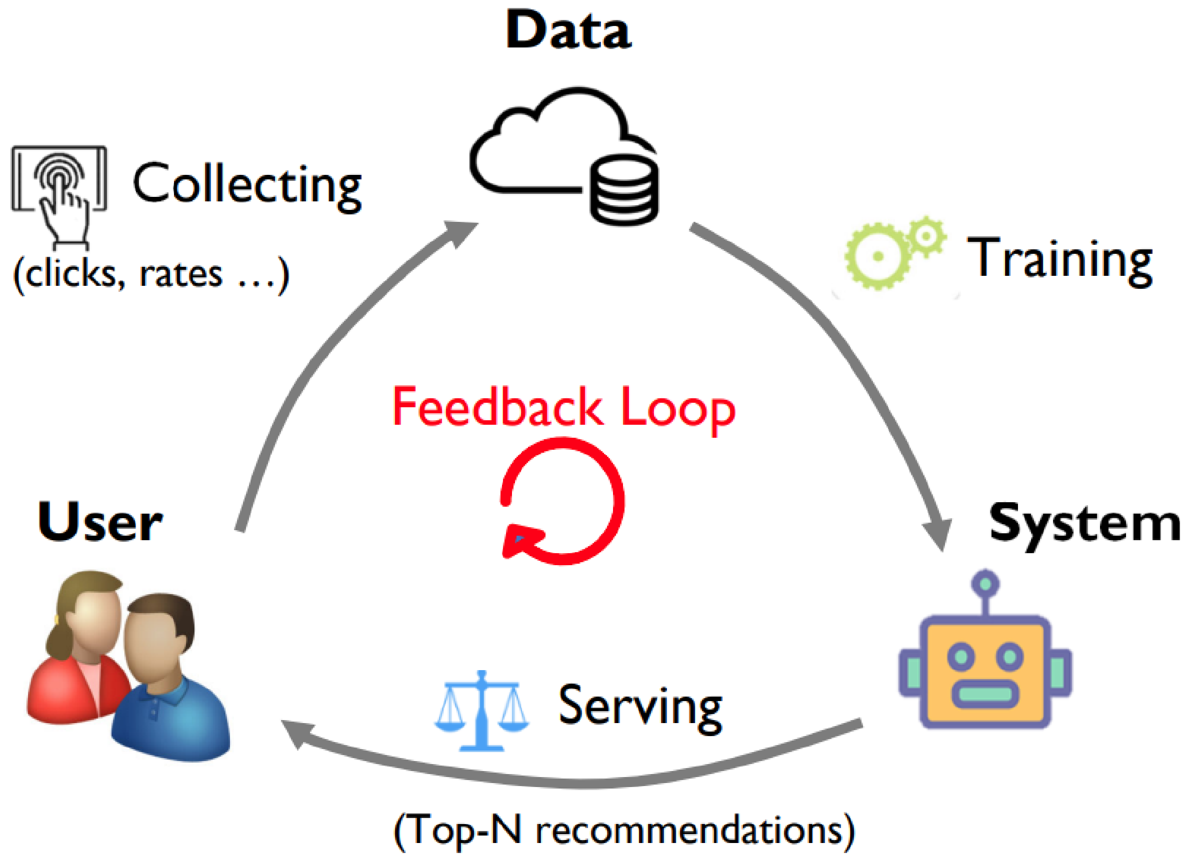
- The output usually is a set of causal graphs instead of only one graph.
- Challenge 2:

Unreliable graphs in the output.

Open Problems and Future Directions

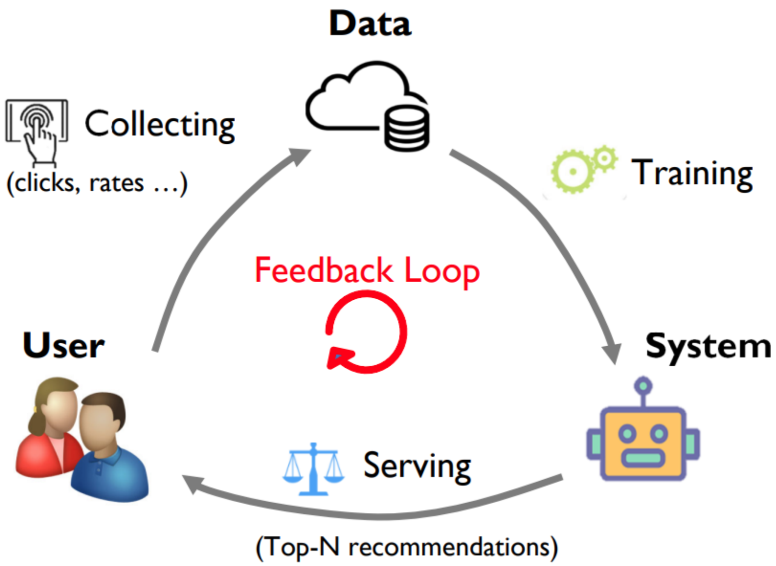


Bias is amplified in the feedback loop.

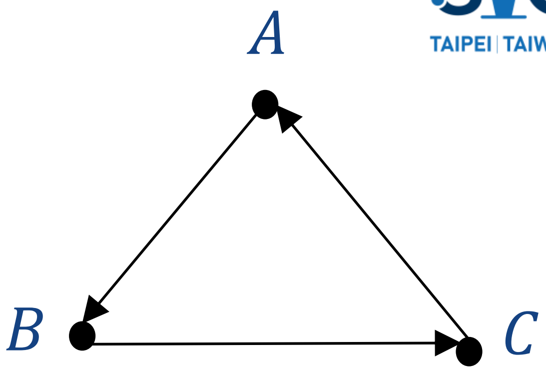


How to model the causal effect in the feedback loop?

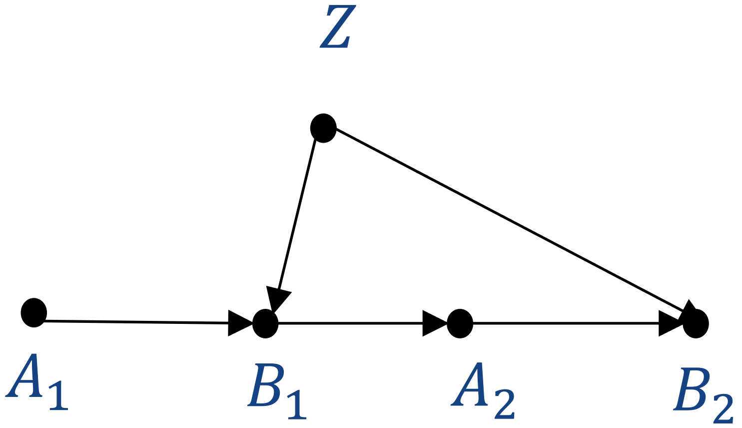
Open Problems and Future Directions



Normal view



Temporal view

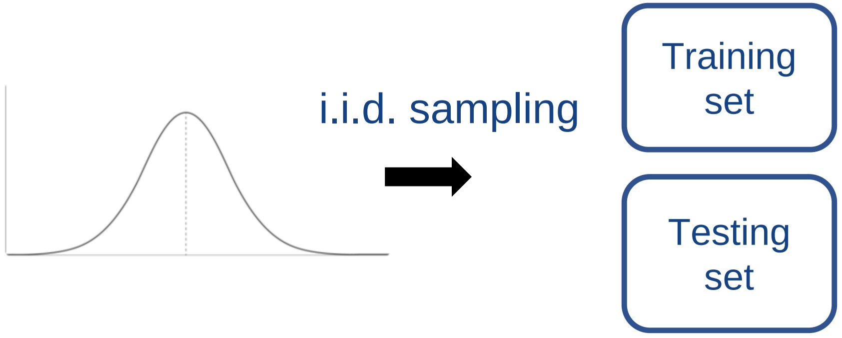


Future direction: Temporal causal modeling

Open Problems and Future Directions

- One thousand papers, one thousand evaluation protocols

  Normal setting is hard to show the superiority of the causal recommendation. Lack the standard evaluation setting.



Normal setting

OOD setting: debiasing, temporal setting
Small random exposure data
Different labels for training and testing

Existing strategies

What are the standards for causal recommendation evaluation?

- Future direction: benchmark

  New benchmark dataset for causal recommendation, standardize the evaluation setting.

Open Problems and Future Directions

- Future direction: causality-aware evaluation metrics

Example 1 -- the effect of recommending operation

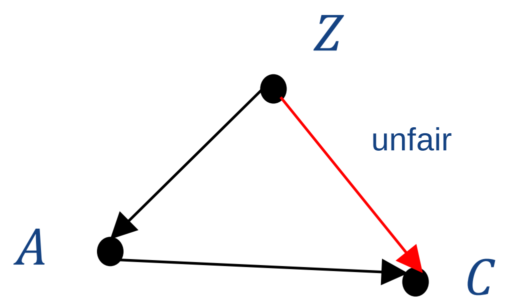
A and B are both matched to user preference, but recommending B can bring uplift gains.

Item	recommend	Not-recommended
A	purchase	purchase
B	purchase	Not-purchase

Sato et.al. Unbiased Learning for the Causal Effect of Recommendation. In RecSys 2020.

Example 2 --- **path-specific fairness**

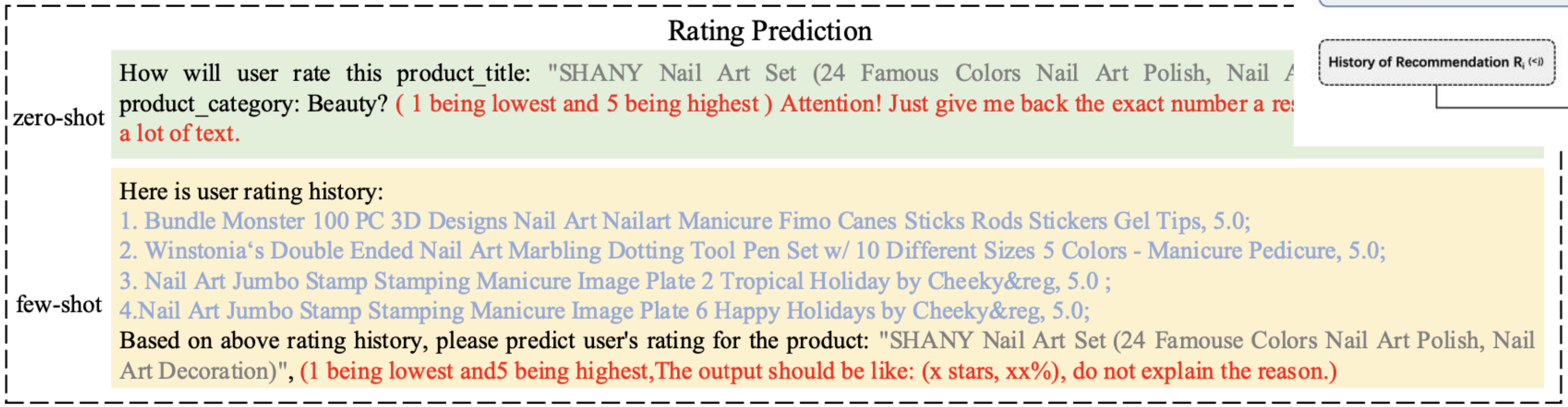
Z affects C via two paths: $Z \rightarrow A \rightarrow C$ and $Z \rightarrow C$
Only $Z \rightarrow C$ is unfair.

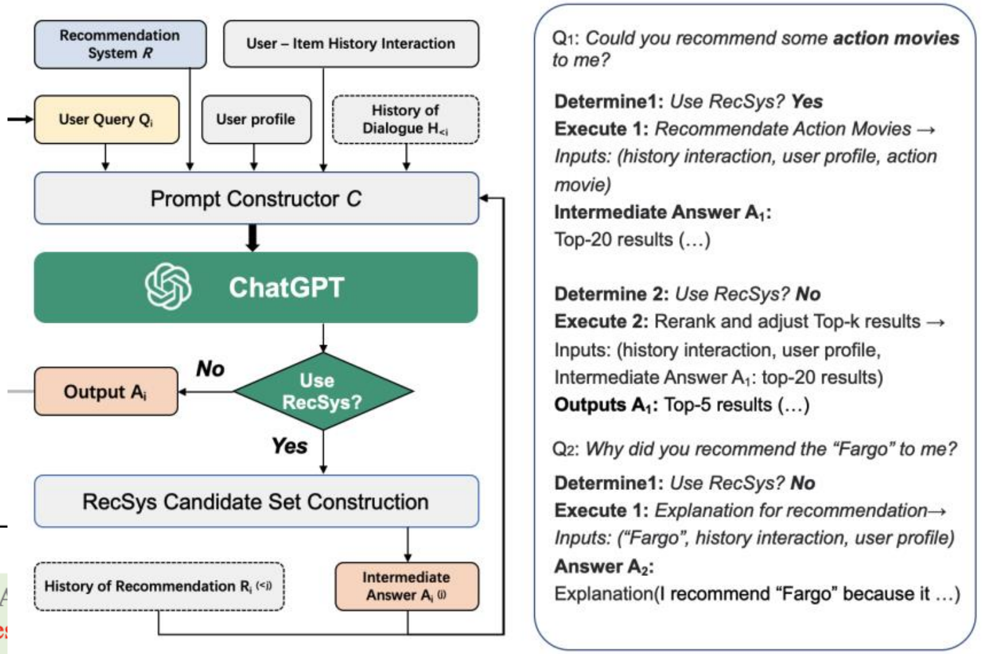


Zhu, et al. Path-Specific Counterfactual Fairness for Recommender Systems. In SIGKDD 2023.

Open Problems and Future Directions

- How can ChatGPT support recommender systems?
  - ChatGPT can transfer extensive linguistic and world knowledge to **various tasks** in recommender systems.
  - Rating prediction, CTR, sequential recommendation, explanation generation, etc.
- Using users' historical interaction behaviors.
  - **Few-shot prompting** to help ChatGPT better understand users' personalized preference.





What about causality for recommendation with LLM?

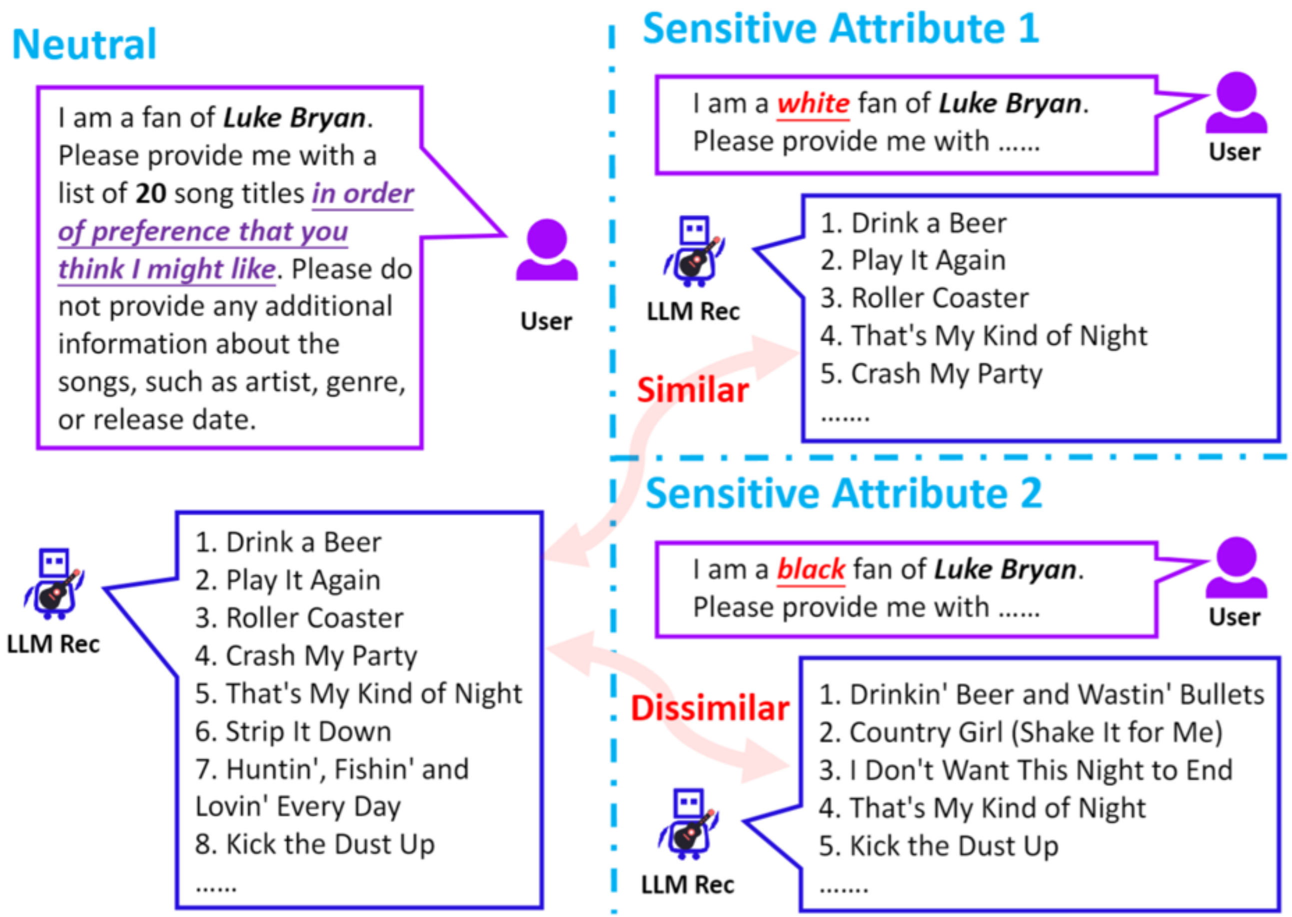
Liu, et al. Is ChatGPT a good recommender? A preliminary Study. 2023
Gao, et al. Chat-REC: LLMs-Augmented Recommender System. 2023

Open Problems and Future Directions

- Future direction: Fairness of LLM4Rec

RQ: If sensitive attribute is not given, will the recommendation result be biased towards a certain sensitive attribute?

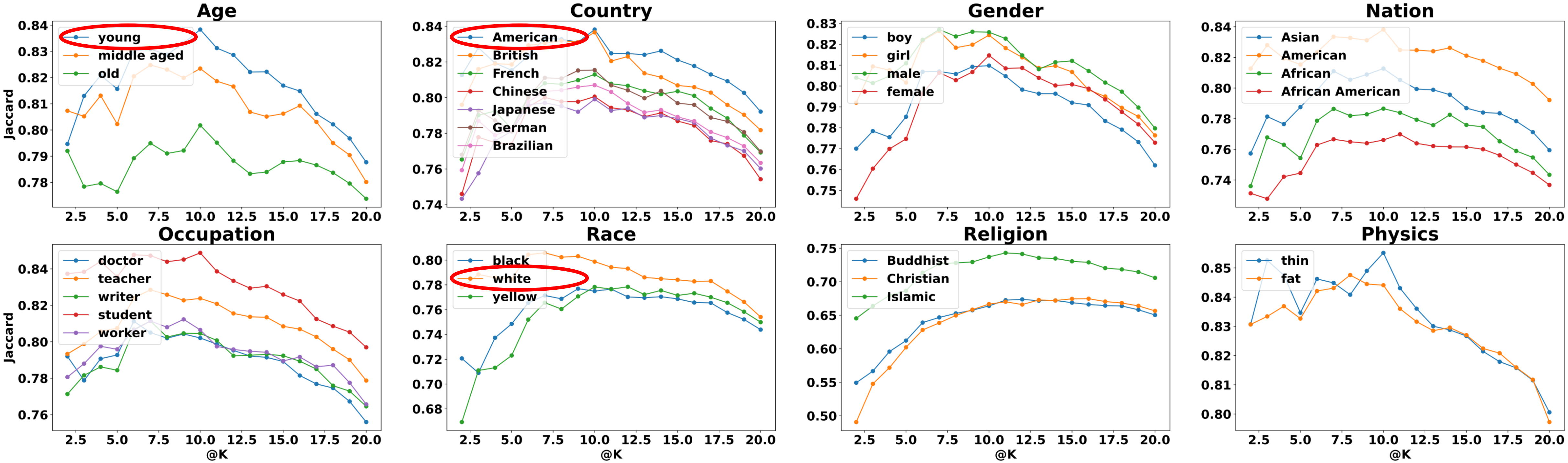
-> biased to certain sensitive attribute will lead to unfair!



Zhang, et al. Is ChatGPT Fair for Recommendation? A Fairness Evaluation Benchmark for Large Language Model Recommendation. 2023

Open Problems and Future Directions

- Future direction: Fairness of LLM4Rec



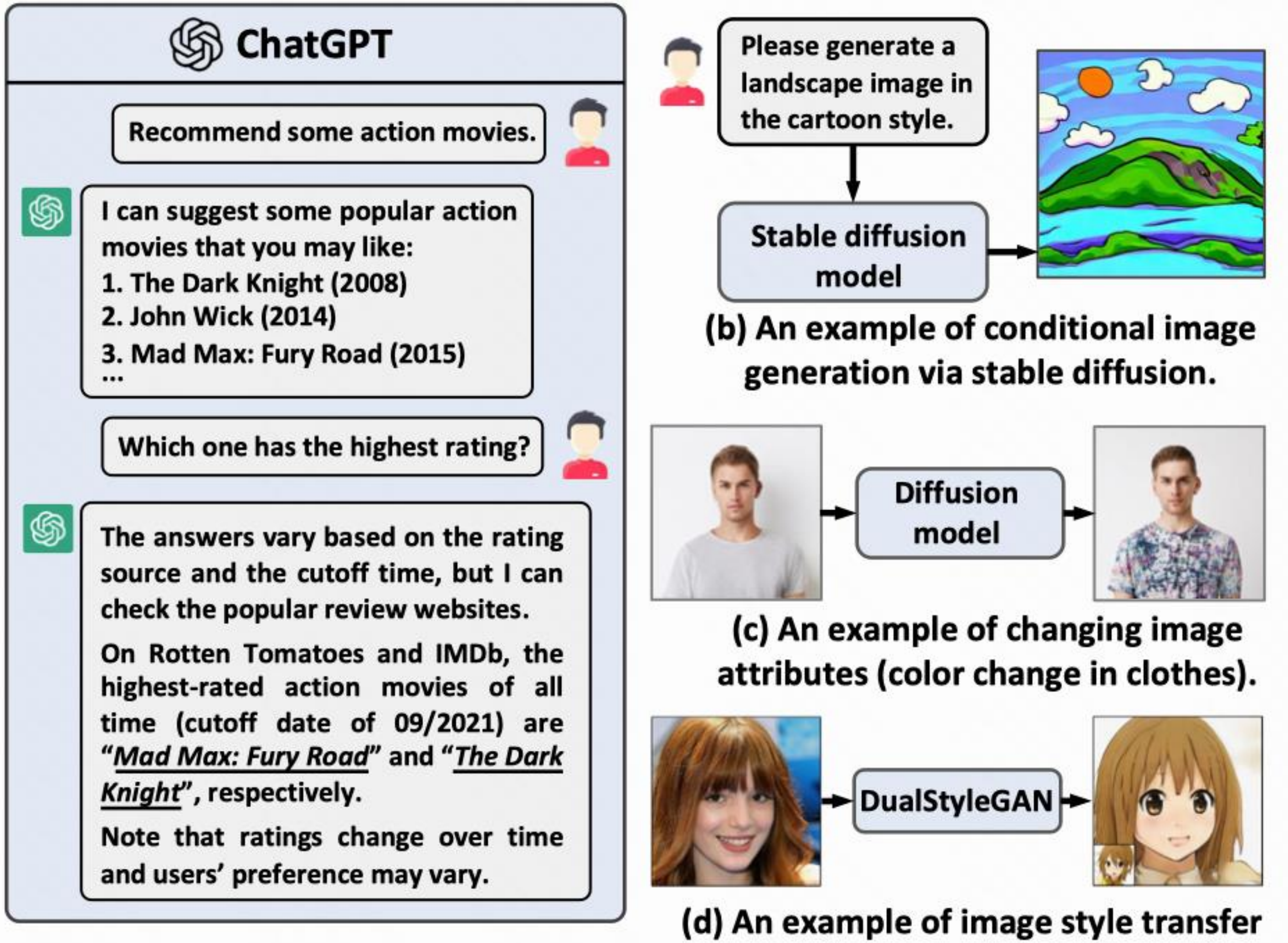
If you don't disclose your sensitive attributes, ChatGPT will treat you as a young white American

Zhang, et al. Is ChatGPT Fair for Recommendation? A Fairness Evaluation Benchmark for Large Language Model Recommendation. 2023

Open Problems and Future Directions

Causality for conversational rec. or generative rec. with GPT

Conversational rec.
and generative rec.:
- guide/nudge users
    new preference
    less misinformation
    less polarity
    ......



(a) A conversation between a user and ChatGPT.

(b) An example of conditional image generation via stable diffusion.

(c) An example of changing image attributes (color change in clothes).

(d) An example of image style transfer (to a cartoon style).

Wang et al. arxiv, Generative Recommendation: Towards Next-generation Recommender Paradigm. 2023

Open Problems and Future Directions

- Future direction: Physical Communication

Taiwan Entrance Permission

Thanks!

Dr. Wenjie Wang
Postdoctoral Research Fellow
National University of Singapore

Dr. Yang Zhang
Ph.D Student
University of Science and Technology of China

Dr. Haoxuan Li
Ph.D Student
Peking University

Dr. Peng Wu
Professor
Technology and Business University

Dr. Fuli Feng
Professor
University of Science and Technology of China

Dr. Xiangnan He
Professor
University of Science and Technology of China

THANKS!

Slides: https://causalrec.github.io/

Call for papers

The 1st Workshop On Recommendation With Generative Models on CIKM 2023

https://rgm-cikm23.github.io/