Estimating Propensity for Causality-based Recommendation without Exposure Data

Zhongzhou Liu[1], Yuan Fang[1], Min Wu[2]
Singapore Management University[1], A*STAR[2]

SMU SINGAPORE MANAGEMENT UNIVERSITY

Background & Motivation

Causal effect of recommendation

	Shampoo	Nintendo Switch
Recommended $Y_{u,i}^1$	95%	50%
Not recommended $Y_{u,i}^0$	90%	10%
Causal effect $\tau_{u,i}$	5%	40%

Causal effect: $\tau_{u,i} = Y_{u,i}^1 - Y_{u,i}^0$

If an item already has a high probability of being interacted by a user without being recommended, ***is there really a need to recommend the item to this user?***

Whether an item has been recommended to a user or not
$ The probability of recommending/exposing an item to a user

- Traditional RS: award item with **higher interaction probability**
- Causal RS: award items with **higher causal effect.**

Motivation

- Existing causal RS assume that *exposure data*# or *propensity scores*$ are observable. **In most public RS datasets, they are not.**
- Other propensity score estimation methods:
  - Require exposure data to train propensity estimator, or
  - Lack prior knowledge (e.g., popularity), less robust estimation.

Preliminaries & Core assumption & Method

Observed variables

- Interactions: $Y_{u,i} \in \{0,1\}$
- Popularity of item i: pop_i

Unobserved variables

- Exposure: $Z_{u,i}$
- Propensity: $p_{u,i} = P(Z_{u,i} = 1)$
- Interaction (**not** exposed): $Y_{u,i}^0$
- Interaction (exposed): $Y_{u,i}^1$

Causal effect estimation [2]

$$\hat{\tau}_{u,i} = \frac{Z_{u,i}Y_{u,i}}{p_{u,i}} - \frac{(1-Z_{u,i})Y_{u,i}}{1-p_{u,i}}$$

Interaction model

$$y_{u,i} = p_{u,i} r_{u,i}$$

Core Assumption

Consider a user u and a pair of items (i, j). Suppose the popularity of item i is greater than that of j, and their interaction probabilities with user u are similar. Then it follows that item i is more likely to be exposed to user u than item j is.■

Overall loss for propensity learning

$$\min_{\Theta} \mathcal{L} = \textstyle\sum_{u,i,j} (\mathcal{L}_{\text{naïve}} + \lambda \mathcal{L}_{\text{pop}}) + \mu \text{KL}(Q \| \text{Beta}(\alpha, \beta))$$

Estimating propensity and exposure

$$\hat{p}_{u,i} = f_p(\mathbf{x}_{u,i}), \quad \hat{Z}_{u,i} = \text{Norm}(\hat{p}_{u,i}) > \epsilon$$

Integrating the assumption into propensity score modelling

$$\mathcal{L}_{\text{pop}} = -\kappa_{u,i,j} \log\left[\sigma(\text{sgn}_{i,j} \cdot (f_p(\mathbf{x}_{u,i}) - f_p(\mathbf{x}_{u,j}))) + \sigma(\text{sgn}_{i,j} \cdot (f_r(\mathbf{x}_{u,j}) - f_r(\mathbf{x}_{u,i})))\right]$$

Causality-based recommendation [2]

$$\frac{Z_{u,i}Y_{u,i}}{\max(p_{u,i}, \chi^1)} \log\left(1 + e^{-\omega(\hat{s}_{u,i} - \hat{s}_{u,j})}\right) + \frac{(1-Z_{u,i})Y_{u,i}}{\max(1-p_{u,i}, \chi^0)} \log\left(1 + e^{\omega(\hat{s}_{u,i} - \hat{s}_{u,j})}\right)$$

Experiment & Conclusion

* PropCare requires no propensity or exposure data in training/reference. They are used only for evaluation.

Methods	DH_original			DH_personalized			ML		
	CP@10↑	CP@100↑	CDCG↑	CP@10↑	CP@100↑	CDCG↑	CP@10↑	CP@100↑	CDCG↑
Ground-truth	.0658±.001	.0215±.001	1.068±.000	.1304±.001	.0445±.001	1.469±.003	.2471±.001	.1887±.000	16.29±.006
Random	.0154±.001	.0071±.002	.7390±.004	.0479±.004	.0107±.005	.8316±.039	.0124±.002	.0135±.005	13.16±.076
POP	.0200±.000	.0113±.000	.7877±.001	.0457±.000	.0096±.001	.8491±.002	-.142±.001	-.092±.001	11.43±.005
CJBPR	.0263±.001	.0087±.001	.7769±.002	.0564±.008	.0106±.005	.8528±.032	-.410±.002	-.187±.001	9.953±.006
EM	.0118±.001	.0067±.001	.7247±.001	.0507±.002	.0121±.001	.8779±.003	-.437±.002	-.194±.002	10.21±.011
PROPCARE	**.0351**±.002	**.0156**±.001	**.9268**±.005	**.1270**±.001	**.0381**±.000	**1.426**±.001	**.0182**±.002	**.0337**±.002	**13.80**±.011

*Results are reported as the average of 5 runs (mean±std). Best results except Ground-truth are bolded, and runners-up are underlined.

PropCare: Drops only 6.6% from ground-truth;
Outperforms other baselines w.r.t. causal recommendation.

Conclusion

- PropCare enables more generalizability of causal RS **without accessing ground-truth exposure & propensity**.
- Core assumption that **integrates prior knowledge** can improve causal RS.

ACKNOWLEDGEMENT & REFERENCES

[1] Xu Xie, et al. CausCF: Causal collaborative filtering for recommendation effect estimation. CIKM 2021.
[2] Masahiro Sato, et al. Unbiased learning for the causal effect of recommendation. RecSys 2020.

For complete references please refer to https://arxiv.org/abs/2310.20388

SCHOOL OF COMPUTING AND INFORMATION SYSTEMS