Counterfactual learning for recommender system

Zhenhua Dong
Hong Zhu
Pengxiang Cheng
Xinhua Feng
Guohao Cai
Xiuqiang He
dongzhenhua@huawei.com
zhuhong8@huawei.com
chengpengxiang1@huawei.com
fengxinhua1@huawei.com
caiguohao1@huawei.com
hexiuqiang1@huawei.com
Huawei Noah's Ark Lab
Shenzhen, China

Jun Xu
Jirong Wen
junxu@ruc.edu.cn
jrwen@ruc.edu.cn
Gaoling School of Artificial Intelligence, Renmin
University of China
Beijing, China

ABSTRACT

Most commercial industrial recommender systems have built their closed feedback loops. Though it is helpful in item recommendation and model training, the closed feedback loop may lead to the so-called bias problems, including the position bias, selection bias and popularity bias. The recommendation models trained with biased may hurt the user experiences by recommending homogenous items. How to control the biases in the closed feedback loop has become one of major challenges in modern recommender systems. This talk discusses the counterfactual learning technologies for tackling the bias problem in recommendation.

The talk consists of four parts.

The first part, briefly introduces the counterfactual learning with two cases from the academic perspective [4, 5].

The second part illustrates the position bias and selection bias based on two real examples. These examples inspire us to study "How to use counterfactual technology for recommender system?" from the industry perspective.

In the third part, we firstly encourage the audiences to think an important question: "What kind of data can learn an unbiased model?" After that, we propose four counterfactual learning approaches and related studies, as shown in Figure1.

Approach 1: Learn from counterfactual data. We need to learn full-information model with partial observed information data. The full-information model is an unbiased model, which is trained by both observed data and unobserved data (including counterfactual data), but how to model unobserved data? One common approach is direct method [2]. In this talk, we introduce a novel counterfactual learning framework [8], first, an imputation model
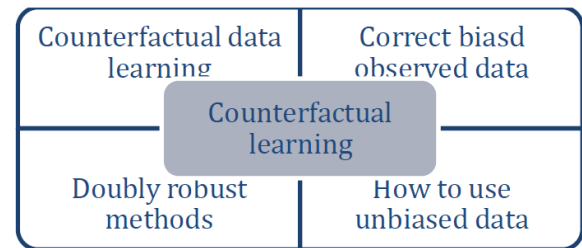
Figure 1: The four counterfactual learning approaches for recommender system

can by learned by a small amount of unbiased uniform data, then the imputation model can be used to predict labels of all counterfactual samples, finally, we train a counterfactual recommendation model with both observed and counterfactual samples.

Approach 2: Correct biased observed data. Inverse propensity score (IPS) is a widely studied method and relatively easy to be deployed for real products. IPS is defined as the conditional probability of receiving the treatment given pre-treatment covariates by Rosenbaum and Rubin [7]. But IPS method should satisfy two assumptions: (1) overlap, and (2) unconfoundedness. Inspired by the sample reweighting work for robust deep learning [6], we proposed a novel influence function based method to reweight training samples directly.

Approach 3: Doubly robust method. The doubly robust methods [7] have two parts: IPS method part and direct method part. John Langford etc. prove that either one part of them can debias, the doubly robust method can debias. But both of the propensity and imputation model are not easy to learn, so we present a novel propensity free doubly robust method [8] for click-through-rate (CTR) prediction task. In order to solve the efficiency of full samples (including both unobserved and observed sample) learning problem, we proposed block coordinate descend and conjugate gradient method, which can reduce the time complexity of optimization from $O(m*n)$ to $O(m+n)$.

Approach 4: Joint learning unbiased data and biased data. In recommender system, unbiased data is collected through randomly recommendation approach. The unbiased data is less, and its collection process is expensive. Through online A/B testing, the performance of the model trained with biased data and unbiased data together is superior to the performance of the model trained with only biased data. Causal embedding [1] method is another method to learn both biased data and unbiased data for improving the accuracy of prediction model. We also propose a general knowledge distillation framework for counterfactual recommendation via uniform data [3], which propose a general framework about how to use unbiased data with four distillation methods: label distillation, sample distillation, feature distillation and model structure distillation.

We also summarize the advantages and challenges of the above approaches.

The last part emphasizes that counterfactual learning is a rich research area, and discuss several important research topics, such as optimization for counterfactual learning, counterfactual meta learning, stable learning, fairness, unbiased learning to rank, offline policy evaluation.

CCS CONCEPTS

• **Information systems** → **Recommender systems**; • **Theory of computation** → **Machine learning theory**.

KEYWORDS

Counterfactual learning, Recommender systems, Causality

ACKNOWLEDGMENTS

REFERENCES

[1] Stephen Bonner and Flavian Vasile. 2018. Causal embeddings for recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems*. 104–112.
[2] Miroslav Dudík, John Langford, and Lihong Li. 2011. Doubly robust policy evaluation and learning. *arXiv preprint arXiv:1103.4601* (2011).
[3] Liu Dugang, Cheng Pengxiang, Dong Zhenhua, He Xiuqiang, Pan Weike, and Zhong Ming. 2020. A general knowledge distillation framework for counterfactual recommendation via uniform data. In *ACM Conference on SIGIR*.
[4] Thorsten Joachims and Adith Swaminathan. 2016. Counterfactual evaluation and learning for search, recommendation and ad placement. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 1199–1201.
[5] Judea Pearl and Dana Mackenzie. 2018. *The Book of Why*. Basic Books, New York.
[6] Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. 2018. Learning to reweight examples for robust deep learning. *arXiv preprint arXiv:1803.09050* (2018).
[7] Paul R Rosenbaum and Donald B Rubin. 1984. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association* 79, 387 (1984), 516–524.
[8] Bowen Yuan, Jui-Yang Hsia, Meng-Yuan Yang, Hong Zhu, Chih-Yao Chang, Zhenhua Dong, and Chih-Jen Lin. 2019. Improving ad click prediction by considering non-displayed events. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. 329–338.