Multiple Robust Learning for Recommendation

Haoxuan Li[1*]**, Quanyu Dai**[2*]**, Yuru Li**[2]**, Yan Lyu**[1]**, Zhenhua Dong**[2]**, Xiao-Hua Zhou**[1]**, Peng Wu**[3,1†]

[1]Peking University
[2]Huawei Noah's Ark Lab
[3]Beijing Technology and Business University
{hxli, lyuyan}@stu.pku.edu.cn, {daiquanyu, liyuru3, dongzhenhua}@huawei.com,
azhou@bicmr.pku.edu.cn, pengwu@btbu.edu.cn

Abstract

In recommender systems, a common problem is the presence of various biases in the collected data, which deteriorates the generalization ability of the recommendation models and leads to inaccurate predictions. Doubly robust (DR) learning has been studied in many tasks in RS, with the advantage that unbiased learning can be achieved when either a single imputation or a single propensity model is accurate. In this paper, we propose 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 a single propensity model, leading to a smaller bias. Inspired by the derived generalization error bound of MR, we further propose a novel multiple robust learning approach for stabilization. We conduct extensive experiments on real-world and semi-synthetic datasets, which demonstrates the superiority of the proposed approach over the state-of-the-art methods.

Introduction

Recommender systems (RSs) have made remarkable progress on many tasks in recent years, such as rating prediction with explicit feedback (Schnabel et al. 2016; Wang et al. 2019), user-item interaction prediction with implicit feedback (Yang et al. 2018; Saito 2019; Saito et al. 2020), post-view click-through rate prediction (Yuan et al. 2019), post-click conversion rate prediction (Zhang et al. 2020; Guo et al. 2021), and uplift modeling (Sato et al. 2019, 2020). Among them, an important challenge is that the collected data are always not the representative of the overall population of interest, due to the presence of selection bias and confounding bias (Chen et al. 2020; Wu et al. 2022b). To address this problem, recent debias-related studies design causality-inspired approaches to achieve unbiased learning for RS. Specifically, Marlin et al. (2012) and Steck (2013) proposed error imputation-based (EIB) approaches, which leverage an additional imputation

model to estimate the prediction errors for missing data. However, the imputed errors tend to be inaccurate in practice and will probably cause a large bias. Besides, Schnabel et al. (2016) proposed the inverse propensity score (IPS) and self-normalized inverse propensity score (SNIPS) methods for unbiased learning with the assistance of a propensity model. Furthermore, Wang et al. (2019) proposed a doubly robust (DR) estimator to combine an error imputation model and a propensity model. DR has double robustness, which is unbiased if either the imputed errors or propensities are accurate. Currently, DR and its enhanced versions (Guo et al. 2021; Dai et al. 2022; Ding et al. 2022) achieve the state-of-the-art performance for debiasing recommendation tasks.

The strength of doubly robust estimators is attributed to the two chances of achieving unbiased learning, i.e., accurate modeling of imputed errors or learned propensities. However, the correct specifications of these models are demanding. If both of the learned models are mildly misspecified, the bias of the doubly robust learning can be severe as pointed out by existing work (Tan 2007; Wu et al. 2022a). What is worse, it is challenging to accurately learn them based on observed data, even if there are correct model specifications. Firstly, the accurate imputed errors are hard to obtain, since the imputation model is learned with the exposed events while being used for the unexposed events. This can easily cause a large bias if the distributions of the exposed events and unexposed events are significantly different (Dai et al. 2022). Secondly, the collected data always contains many biases (Chen et al. 2020; Wu et al. 2022b) (e.g., biases induced by unobserved confounders), which may result in the inaccurate estimation of propensities. Therefore, more research efforts for unbiased learning are needed.

In this paper, we propose a multiple robust (MR) estimator for unbiased learning in RS, which contains multiple candidate propensity models and error imputation models. It allows multiple different specifications and learning of the propensity and imputation models. Through theoretical analysis, we demonstrate that the proposed MR estimator has multiple robustness, which is unbiased if any of the propensity models, imputation models, or even a linear combination of these models can accurately estimate the true propensities or prediction errors. Therefore, the MR estimator can significantly tackle the problem of inaccuracy in learned propensities or

---

imputed errors from a single model of existing work, by providing more opportunities for model specification and learning. In addition, we particularly analyze the relation between MR and DR, when MR has only one propensity model and one imputation model, and theoretically find that our MR estimator can actually provide an enhanced version of double robustness. By analyzing the tail bound of the MR estimator, we further propose a multiple robust learning approach for stabilization, to better control the generalization error by adding an L_2 regularization. Extensive experiments on both real-world and semi-synthetic datasets demonstrate that the proposed MR method achieves a significant improvement compared to the state-of-the-art methods.

Related Work

Unbiased Learning in Recommendation. RS is always affected by various biases (Wu et al. 2022b), leading to the inconsistency between the collected data and the target population. Therefore, unbiased learning has been extensively studied (Marlin and Zemel 2009; Steck 2010; Liu et al. 2021). In RS, Schnabel et al. (2016) and Saito (2019) used a single propensity model for unbiased learning on explicit and implicit feedback tasks, while Marlin et al. (2012) and Steck (2013) adopted a single error imputation model for unbiased learning. Wang et al. (2019) proposed to combine a propensity and an error imputation model to achieve doubly robust learning, and yielded superior performance on ranking (Saito 2020; Kiyohara et al. 2022), uplift modeling (Sato et al. 2020) and data fusion (Bonner and Vasile 2018; Liu et al. 2020; Chen et al. 2021; Wang et al. 2021; Li et al. 2023b). Based on that, Guo et al. (2021), Dai et al. (2022), Li et al. (2023a) and Li, Zheng, and Wu (2023) proposed more robust DR methods, and Zhang et al. (2020) proposed a parameter sharing mechanism applied to multi-task learning. In addition, Ding et al. (2022) proposed a robust deconfounder framework to mitigate the influence of unmeasured confounder. Different from the existing debiasing methods, this paper develops a novel multiple robust learning approach, which allows multiple propensity and error imputation models and thus enables a greater chance of unbiased learning.

Multiple Robust. MR approaches were designed to prevent the inaccuracy of learned propensities and imputed errors. Currently, the MR methods mainly focused on estimating the population mean of a response variable with missing (Li, Gu, and Liu 2020). They can be divided into two major categories: empirical likelihood-based calibration methods (Han and Wang 2013; Chen and Haziza 2017; Han 2016, 2018) and ensemble learning-based methods (Chan 2013; Chan and Yam 2014; Duan and Yin 2017). The former relies on solving a large optimization problem and is hard to implement in the data of RS. To ease the computational burden, the latter decomposes the optimization problem into two steps, which first compress multiple propensity/imputation models into one new propensity/imputation model, and then solve an optimization problem the same as the calibration approaches (Duan and Yin 2017). How to reduce the computational burden and how to embed them in a joint learning model in RS are challenges problems. To the best of our knowledge, this is the first work that proposes a multiple

robust learning approach for debiased recommendations.

The Proposed Method

Problem Definition and Preliminaries

Let $\mathcal{U} = \{u\}$ be the users, $\mathcal{I} = \{i\}$ the items, and $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ all of the user-item pairs. Let $x_{u,i}$ be the feature of a user-item pair (u, i), which is used by the prediction model $f(x_{u,i}; \theta)$ with model parameters θ in RS, to predict the rating $y_{u,i}$ of item i given by user u. In RS, a large fraction of $y_{u,i}$ is missing-not-at-random, resulting in the highly sparse and biased data along with the challenge of unbiased model learning. Define $o_{u,i}$ as an indicator function of whether $y_{u,i}$ is observed or not, i.e., $o_{u,i} = 1$ if $y_{u,i}$ is observed and $o_{u,i} = 0$ otherwise. Denote $\mathcal{O} = \{(u, i) \in \mathcal{D} : o_{u,i} = 1\}$ as the indices of observed samples, and $\hat{\mathbf{Y}}$ as the predicted rating matrix of all user-item pairs, obtained from the prediction model $f(x_{u,i}; \theta)$. This work aims to train the prediction model based on the observed samples to infer all true ratings $y_{u,i}$ on $(u, i) \in \mathcal{D}$.

To evaluate the performance of the prediction model $f(x_{u,i}; \theta)$ over all user-item pairs, let $e_{u,i}$ be the prediction error, which is an appropriately chosen loss function, e.g., the squared loss $(y_{u,i} - f(x_{u,i}; \theta))^2$. However, the unbiased evaluation on $\mathcal{D}$ requires all values of $y_{u,i}$. Given all the predicted ratings $\hat{y}_{u,i} \triangleq f(x_{u,i}; \theta)$, ideally, if all ratings $y_{u,i}$ are observed, the prediction model $f(\cdot; \theta)$ can be trained by directly optimizing the following ideal loss

$$\mathcal{E}_{ideal} = \mathcal{E}_{ideal}(\hat{\mathbf{Y}}) = |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} e_{u,i}.$$

However, since $y_{u,i}$ is observed only when $o_{u,i} = 1$, the ideal loss is non-computable. To address this problem, many debiasing methods in RS have been proposed by constructing unbiased substitutes of the ideal loss, such as EIB estimator $\mathcal{E}_{EIB} = |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} [o_{u,i} e_{u,i} + (1 - o_{u,i}) \hat{m}_{u,i}]$, and IPS estimator $\mathcal{E}_{IPS} = |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} o_{u,i} e_{u,i} / \hat{p}_{u,i}$, where $\hat{m}_{u,i} \triangleq m(x_{u,i}; \hat{\beta})$ is an estimate of the prediction error $e_{u,i}$ from an error imputation model with parameter $\hat{\beta}$, and $\hat{p}_{u,i} \triangleq \pi(x_{u,i}; \hat{\alpha})$ is an estimate of propensity score $p_{u,i} \triangleq \mathbb{P}(o_{u,i} = 1 | x_{u,i})$ from a propensity model with parameter $\hat{\alpha}$. Besides, the doubly robust (DR) uses both imputed errors and propensities to further reduce the bias, which is

$$\mathcal{E}_{DR} = |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} \left[\hat{m}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{m}_{u,i})}{\hat{p}_{u,i}} \right].$$

Multiple Robust Estimator

The unbiasedness of IPS and EIB estimators rely on the accurate estimations of propensities and prediction errors, respectively, from a single model. The DR estimator takes advantage of both the estimated propensities and imputed errors to achieve double robustness, but it only allows to use one propensity model and one imputation model. Nevertheless, the challenges of model specification and model learning significantly hinder the realization of unbiased learning. In this work, we propose a multiple robust (MR) estimator to

tackle these challenges by introducing multiple propensity models and error imputation models. Specifically, 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 $\hat{p}_{u,i}^j \triangleq \pi_j(x_{u,i}; \hat{\alpha}_j)$, $\hat{m}_{u,i}^k \triangleq m_k(x; \hat{\beta}_k)$, and

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

The proposed MR estimator is given as

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

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

$$\sum_{(u,i) \in \mathcal{D}} o_{u,i} \{e_{u,i} - \boldsymbol{u}^T(x_{u,i}) \cdot \boldsymbol{\eta}\}^2. \quad (1)$$

Denote $\hat{\boldsymbol{\eta}}(\theta)$ as $\hat{\boldsymbol{\eta}}$ for simplification. Let $[\hat{\mathbf{P}}^j]_{u,i} = 1/\hat{p}_{u,i}^j$ be the (u, i)-th element of the j-th learned inverse propensities matrix, and $[\hat{\mathbf{E}}^k]_{u,i} = \hat{m}_{u,i}^k$ be the (u, i)-th element of the k-th imputed errors matrix.

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.

We formally elaborate the multiple robustness of the MR estimator. Firstly, note that $\hat{\boldsymbol{\eta}}$ will make the first-derivative of equation (1) with respect to $\boldsymbol{\eta}$ equals to 0, i.e.,

$$\sum_{(u,i) \in \mathcal{D}} o_{u,i} \cdot [e_{u,i} - \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}] \cdot \boldsymbol{u}(x_{u,i}) = 0^{(J+K) \times 1},$$

of which the j-th element is that

$$\sum_{(u,i) \in \mathcal{D}} \{e_{u,i} - \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}\} o_{u,i}/\hat{p}_{u,i}^j = 0.$$

Therefore, for any given $j = 1, ..., J$,

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

$$\underbrace{\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}^j} \cdot \{e_{u,i} - \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}\}}_{0}$$

$$= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}^j} \cdot e_{u,i} +$$

$$\frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[1 - \frac{o_{u,i}}{\hat{p}_{u,i}^j}\right] \cdot \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}. \quad (2)$$

---

[1]An estimator is consistent if, as the sample size increases, it converges to the true value of the parameter in probability.

Suppose $1/\hat{p}_{u,i}^j$ is an accurate estimate of $1/p_{u,i}$ and it can be shown that the last term in equation (2) will converge to 0 in probability, which implies that

$$\mathbb{E}_{\mathcal{O}}[\mathcal{E}_{MR}] \xrightarrow{P} |\mathcal{D}|^{-1} \sum_{(u,i) \in \mathcal{D}} \mathbb{E}_{\mathcal{O}}[o_{u,i} e_{u,i}/\hat{p}_{u,i}^j] = \mathcal{E}_{ideal}.$$

If $[\hat{\mathbf{P}}^{ln}]_{u,i} = 1/p_{u,i}$ for some weights $w_1, ..., w_J$, the consistency of $\mathcal{E}_{MR}$ follows immediately from the weighted summation of the equation (2) with different j.

Instead, if one of the imputed errors, or a linear combination of the K imputed errors can estimate $e_{u,i}$ accurately, i.e., $[\hat{\mathbf{E}}^{ln}]_{u,i} = e_{u,i}$ for some weights $v_1, ..., v_K$, which implies that $\hat{\boldsymbol{\eta}} = (0, \cdots, 0, v_1, ..., v_K)^T$, resulting in $\mathbb{E}_{\mathcal{O}}[\mathcal{E}_{MR}] = \mathcal{E}_{ideal}$. Thus, $\mathcal{E}_{MR}$ is multiple robust in the sense of consistency. In addition, when using different samples to calculate $\hat{\boldsymbol{\eta}}$ and $\mathcal{E}_{MR}$, such independence ensures that the expectation of last term in equation (2) is strictly equal to 0, resulting in $\mathbb{E}_{\mathcal{O}}[\mathcal{E}_{MR}] = \mathcal{E}_{ideal}$ (See the next section for implementation details).

Theorem 1 indicates that the MR estimator can achieve unbiasedness under weaker conditions than those of the DR estimator. Instead of IPS and EIB relying only on a single model, and DR relying on a single propensity and a single imputation model, the MR estimator allows to exploit multiple candidate models to reach unbiasedness. In addition, MR has great potential in industrial application scenarios of RS, since there are usually many candidate models for recommendation tasks, and MR further allows combining those models as shown in Theorem 1 and enhances the prediction accuracy.

When the MR estimator contains exactly *one error imputation model* $\hat{m}_{u,i}$ and *one propensity model* $\hat{p}_{u,i}$, a natural question is the relationship between MR and DR, and Theorem 2 states that MR acts as a boosted version of DR.

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.

Note that DR can be written as

$$\mathcal{E}_{DR} = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[\frac{o_{u,i} \cdot e_{u,i}}{\hat{p}_{u,i}} + \left\{1 - \frac{o_{u,i}}{\hat{p}_{u,i}}\right\} \cdot \hat{m}_{u,i}\right], \quad (3)$$

and it follows from equation (2) that

$$\mathcal{E}_{MR} = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left[\frac{o_{u,i} \cdot e_{u,i}}{\hat{p}_{u,i}} + \left\{1 - \frac{o_{u,i}}{\hat{p}_{u,i}}\right\} \cdot \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}\right]. \quad (4)$$

Therefore, $\mathcal{E}_{MR}$ degenerates to $\mathcal{E}_{DR}$ if $\hat{m}_{u,i} = \boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}$, for which a sufficient condition is $e_{u,i} = \hat{m}_{u,i}$, i.e., the error imputation model is accurate.

By a direct comparison of equation (3) and equation (4), it can be seen that MR performs as a different estimator

with double robustness, replacing the $\hat{m}_{u,i}$ with $\boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}$. Importantly, if $1/\hat{p}_{u,i}$ is helpful to fit $e_{u,i}$, then $\boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}$ is expected to be a more accurate estimate of the prediction error than $\hat{m}_{u,i}$. If $1/\hat{p}_{u,i}$ doesn't help, then $\boldsymbol{u}^T(x_{u,i}) \cdot \hat{\boldsymbol{\eta}}$ and $\hat{m}_{u,i}$ will have the same accuracy. Thus, $\mathcal{E}_{MR}$ can be viewed as an improved version of DR when we only have access to an error imputation model and a propensity model.

Theorem 1 shows that the bias of MR vanishes with an accurate model. However, when all of the models are inaccurate, the bias of MR is derived in Theorem 3.

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

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

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

Theorem 3 shows that the bias of MR estimator consists of a dominant term and a negligible term of order $O(|\mathcal{D}|^{-1})$. In addition, it can be seen that the bias of MR has similar form compared to that of DR, based on the j-th propensity model and k-th imputation model, which is given by

$$\frac{1}{|\mathcal{D}|}\left| \sum_{(u,i)\in\mathcal{D}} \underbrace{\left\{1 - \frac{p_{u,i}}{\hat{p}_{u,i}^j}\right\}}_{\text{single propensity } 1/\pi_j} \cdot \underbrace{\left\{e_{u,i} - \hat{m}_{u,i}^k\right\}}_{\text{single imputation } m_k} \right|.$$

Note that both the biases consist of the product of the estimation inaccuracies from learned propensities and imputed errors. However, MR allows using the optimal linear combination of multiple propensity models to approximate $p_{u,i}$, as well as using the optimal linear combination of all models to approximate $e_{u,i}$, while DR only uses one single propensity and imputation model to approximate $p_{u,i}$ and $e_{u,i}$, respectively. This shows that MR tends to have a smaller bias compared to that of the DR.

Alternating Multiple Robust Learning with Stabilization

In this subsection, we analyze the tail bound of the MR estimator and the generalization error bound of the learned prediction model based on the MR loss, and find that both of them depend on the norm of $\hat{\boldsymbol{\eta}}$, which motivates us to propose an alternative multiple robust learning for stabilization. The following Theorem 4 presents the tail bound of the MR.

Theorem 4 (Tail Bound of MR). *Suppose all J propensity models satisfy $\Gamma^{-1} \leq \hat{p}_{u,i}^j \leq 1$ $(\Gamma \geq 1)$ for $j = 1,\ldots,J$, and all K imputation models satisfy $|\hat{m}_{u,i}^k| \leq M/2$ for $k = 1,\ldots,K$. Then for any prediction matrix $\hat{\mathbf{Y}}$ with given*

$\hat{\boldsymbol{\eta}}$*, with probability $1 - \delta$, the MR estimator does not deviate from its expectation by more than*

$$\left| \mathcal{E}_{MR}(\hat{\mathbf{Y}}) - \mathbb{E}_O\left[\mathcal{E}_{MR}(\hat{\mathbf{Y}})\right] \right| \leq \sqrt{\frac{\log(2/\delta)}{2|\mathcal{D}|}}\max(\Gamma - 1, M) \cdot \|\hat{\boldsymbol{\eta}}\|_1.$$

For any given hypothesis space $\mathcal{H}$ of the prediction model, the optimal prediction model $\hat{\mathbf{Y}}^\dagger$ can be defined by $\hat{\mathbf{Y}}^\dagger = \arg\min_{\hat{\mathbf{Y}}\in\mathcal{H}} \mathcal{E}_{MR}(\hat{\mathbf{Y}})$. Then, the generalization error bound of $\hat{\mathbf{Y}}^\dagger$ from $\mathcal{H}$ obtained using MR can be further derived as shown in Theorem 5.

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*

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

Theorem 5 shows that the generalization error bound can be controlled via the tighter range bounds of the propensity and imputation models, as well as a smaller value of $\|\hat{\boldsymbol{\eta}}\|_1$. The former can be achieved by clipping the extremely small estimated propensities or imposing an additional penalty for extreme imputed errors. The latter can be achieved by imposing an additional regularization term on the MR loss to penalize the norm of $\hat{\boldsymbol{\eta}}$ to stabilize the MR estimator.

A key design of the learning process is to use different samples to compute $\hat{\boldsymbol{\eta}}$ and update the prediction model with MR loss. As shown in Theorem 1, such independence guarantees the strict unbiased learning process. Specifically, we first randomly sample a batch $\mathcal{D}'$ from $\mathcal{D}$, and use the samples $\mathcal{O}'$ with observed ratings in $\mathcal{D}'$ to compute $\hat{\boldsymbol{\eta}}$. Next, another batch is drawn from the remaining samples $\mathcal{D} \setminus \mathcal{D}'$ for the calculation of MR loss and the update of the prediction model.

Since Lasso has no closed-form solution and the iterative procedure will increase the computational cost, we take the Ridge regression with $\mathcal{O}' \subset \mathcal{D}'$ to obtain $\hat{\boldsymbol{\eta}}$ in MR

$$\hat{\boldsymbol{\eta}} = \left[\sum_{(u,i)\in\mathcal{O}'} \boldsymbol{u}(x_{u,i}) \cdot \boldsymbol{u}^T(x_{u,i}) + \lambda\boldsymbol{I}\right]^{-1}\left[\sum_{(u,i)\in\mathcal{O}'} \boldsymbol{u}(x_{u,i}) \cdot e_{u,i}\right], \tag{5}$$

where $\boldsymbol{I}$ is an identity matrix, and λ is a hyper-parameter for stabilization. Then the prediction model is updated by the proposed MR loss $\mathcal{L}_{MR}(\theta; \hat{\alpha}, \hat{\beta}) = \sum_{(u,i)\in\mathcal{D}\setminus\mathcal{D}'} \boldsymbol{u}(x_{u,i})^T\hat{\boldsymbol{\eta}}(\theta)$. In addition, given the prediction model and sampling data $\{(u_{k_l}, i_{k_l})\}_{l=1}^L$ for $k = 1, \ldots, K$ from $\mathcal{O}$, the k-th imputation model is updated by using

$$\mathcal{L}_{e_k}(\theta, \beta_k) = L^{-1}\sum_{l=1}^{L}(m_k(x_{u_{k_l},i_{k_l}}; \beta_k) - e_{u_{k_l},i_{k_l}})^2/\hat{p}_{u_{k_l},i_{k_l}}^j,$$

where $e_{u,i} = y_{u,i} - f_\theta(\boldsymbol{x}_{u,i})$ and $\hat{p}_{u,i}^j$ is a randomly-chosen propensity model. By alternatively updating the imputation models and prediction ratings with the above process, we can finally achieve unbiased learning with stabilization. We summarized the alternating learning approach in Alg. 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

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

Experiments

In this section, we conduct experiments on both real-world datasets and semi-synthetic datasets to evaluate the effectiveness of our proposed method.

Baselines. The proposed MR[2] and most existing debiasing methods are model-agnostic, which can be integrated into existing recommendation models for unbiased learning based on biased data. We follow prior studies (Wang et al. 2020; Liu et al. 2021) to adopt two of the most common recommendation models as the backbones, i.e., matrix factorization (MF) (Koren, Bell, and Volinsky 2009) and neural collaborative filtering (NCF) (He et al. 2017). Besides the naive estimator, we compare our method with the following state-of-the-art debiasing estimators: inverse propensity score (IPS) method (Schnabel et al. 2016), the self-normalized IPS (SNIPS) (Swaminathan and Joachims 2015), doubly robust learning (DR) (Jiang and Li 2016), doubly robust joint learning (DR-JL) (Wang et al. 2019), more robust doubly robust joint learning (MRDR-JL) (Guo et al. 2021), counterfactual variational information bottleneck (CVIB) (Wang et al. 2020), and debiased information bottleneck (DIB) (Liu et al. 2021).

Evaluation Protocols and Experimental Details. We employ three widely used metrics to measure the prediction performance on the testing set for unbiased evaluation, in-

cluding the mean square error (MSE), the area under the ROC curve (AUC), and the normalized discounted cumulative gain (NDCG). MSE measures the accuracy of the prediction, AUC measures the overall ranking performance on the testing set, and NDCG evaluates the ranking performance in a user-wise manner. All experiments are implemented on PyTorch (Paszke et al. 2019) with Adam optimizer (Kingma and Ba 2015), and grid search is used to choose the optimal set of hyper-parameters based on a validation set split from the training set.

Experiments on Real-World Datasets

Datasets. We consider two benchmark real-world datasets containing MNAR and MAR ratings, i.e., **Coat**[3] (Schnabel et al. 2016) and **Yahoo**[4] (Marlin and Zemel 2009), as existing work (Schnabel et al. 2016; Wang et al. 2019). Specifically, Coat has 6,960 five-star ratings from 290 Amazon Mechanical Turkers on an inventory of 300 coats in the training set, and 4,640 ratings collected from the 290 workers on 16 randomly selected coats in the testing set. Yahoo includes a MNAR training set with 311,704 five-star ratings from 15,400 users and 1,000 songs, and a MAR testing set with 54,000 ratings from 5,400 users on 10 randomly selected songs. We follow prior studies (Schnabel et al. 2016; Wang et al. 2019, 2020) to use the MNAR dataset for training and the MAR dataset for unbiased evaluation on both datasets.

Performance. The experimental results are shown in Table 1. We have the following observations. Existing debiased estimators, including IPS, SNIPS, DR, DR-JL, MRDR-JL, CVIB and DIB, can effectively improve recommendation performance under biased training sets, since they all produce better results than the corresponding backbone models. DR methods only perform comparable to IPS based methods, which might be due to the inaccurate imputed errors. Impressively, our proposed MR estimator achieves significant improvements over the existing debiased methods in both Coat and Yahoo regardless of the backbone recommendation models. The reason can be attributed to its multiple robustness. By utilizing multiple different propensity models and imputation models, the MR estimator achieves state-of-the-art debiasing performance in terms of MSE, AUC and NDCG@K.

Experiments on Semi-synthetic Datasets

Datasets. We conduct experiments on semi-synthetic datasets constructed from **MovieLens 100K**[5] (**ML-100K**) to evaluate the robustness of our proposed method. ML-100K is a dataset collected from a movie recommendation service with 100,000 MNAR ratings from 943 users and 1,682 movies. We generate three datasets based on it, denoted as ML-100K-1, ML-100K-2, and ML-100K-3, which have an increasing level of selection bias by varying the skewness of the propensities.

Performance. The experimental results are shown in Table 2. It can be found that our proposed MR estimator achieves the best results among all three semi-synthetic datasets, which

---

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.

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

Datasets	ML-100K-1			ML-100K-2			ML-100K-3		
Methods	MSE	AUC	NDCG@10	MSE	AUC	NDCG@10	MSE	AUC	NDCG@10
Naïve	0.1188	0.6587	0.3223	0.1454	0.5700	0.2452 (-23.9%)	0.2023	0.5105	0.2009 (-37.7%)
IPS	0.0913	0.7260	0.3517	0.1025	0.6321	0.2889 (-17.9%)	0.2043	0.5144	0.2227 (-36.7%)
SNIPS	0.0914	0.7251	0.3526	0.1023	0.6318	0.2867 (-18.7%)	0.1975	0.5126	0.2282 (-35.3%)
DR	0.0997	0.6678	0.3204	0.1021	0.5926	0.2613 (-18.4%)	0.1825	0.5186	0.2075 (-35.2%)
DR-JL	**0.0896**	0.7296	0.3536	0.0975	0.6418	0.2875 (-18.7%)	0.0991	0.5959	0.2838 (-19.7%)
MRDR-JL	0.0904	0.7268	0.3521	0.0980	0.6440	0.2919 (-17.1%)	0.0987	0.6036	0.2823 (-19.8%)
CVIB	0.1036	0.6596	0.3151	0.1040	0.5538	0.2156 (-31.6%)	0.2903	0.5050	0.1786 (-43.3%)
DIB	0.1190	0.6583	0.3243	0.1407	0.5756	0.2531 (-22.0%)	0.1908	0.5117	0.1878 (-42.1%)
MR (Ours)	0.0902	**0.7316**	**0.3622**	**0.0919**	**0.6968**	**0.3318** (-8.4%)	**0.0980**	**0.6379**	**0.3100** (-14.4%)

* The best results are highlighted in bold.

Table 2: Experimental results on the three semi-synthetic datasets, including ML-100K-1, ML-100K-2, and ML-100K-3, which have increasing exposure bias. We also present the relative performance drops of all methods on NDCG@10 in ML-100K-2 and ML-100K-3 compared with the corresponding results in ML-100K-1 to demonstrate the robustness of different models.

demonstrates MR to be the most effective unbiased estimator compared with existing methods. Besides, MR demonstrates the most competitive performance in terms of robustness. Specifically, with the increase of selection bias, the performance of IPS and SNIPS drops dramatically (e.g., the AUC score of IPS drops up to 17.9% and 36.7% in ML-100K-2 and ML-100K-3, respectively, compared with that in ML-100K-1). Both DR-JL and MRDR-JL outperform IPS and SNIPS due to their double robustness. Our proposed MR has superior performance in all three datasets and shows significantly little performance drops with the increase of data bias, because of its multiple robustness for unbiased learning.

In-Depth Analysis of MR

In this subsection, we conduct an in-depth analysis to study the effect of the imputation models and propensity models.

Effect of Imputation Models. We study the effect of im-

putation models on the MR estimator by setting MR with different numbers and types of imputation models. We use a single propensity model, i.e., the Naive Bayesian model, for all experiments for comparison purpose. MF is used as the backbone prediction model, while both MF and NCF are used for imputations depending on specific configurations.

As shown in the left side of Table 3, when increasing the number of imputation models from one to five, all evaluation metrics first improve steadily and significantly, and then slightly drop. The best performance is achieved in the MR with four imputation models. It demonstrates that utilizing multiple imputation models instead of only one in the MR estimator can effectively improve model robustness and generalization. Theoretically, as long as one of the imputation models estimate the prediction error accurately, the MR approach can achieve unbiased learning. However, it is worth noting that the MR method with five imputation models has

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

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

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**

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.
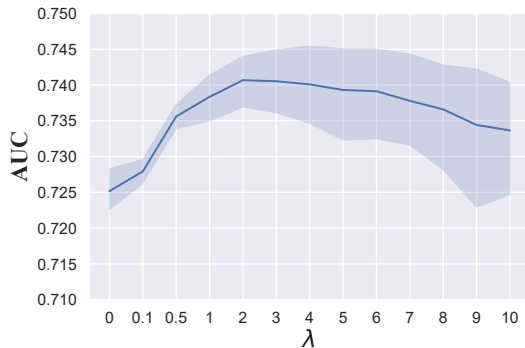

Figure 1: AUC of MR with different settings of λ. The 90% confidence of the results is shown with the shaded regions.

slightly worse performance compared with that of fours. It can be interpreted as the additional MF model brings less improvement to the prediction accuracy than the noise. Therefore, the choice of the number of imputation models can be considered as a bias-noise trade-off, and the proper number of imputation models will lead to the optimal performance.

We further investigate the effect of the combination of different types of imputation models on the performance of MR methods, and the results are shown in the right side of Table 3. First, MR methods with two MF and NCF imputations obtain an absolute increase in AUC scores of 3.4‰ and 5.1‰, respectively, compared to those methods with only one MF and NCF imputation model. Furthermore, although the mixture of MF and NCF performed competitively, the MR with two NCF imputations performs best. This suggests that deep recommendation models such as NCF perform better than simple models such as MF in terms of estimating prediction errors. Overall, the proper number of deep imputation models can help improve the debiasing performance of MR.

Effect of Propensity Models. Table 4 shows the experimental results including different settings of propensity models on MR. Specifically, 'NB', 'NB-Uni', and 'User' represent the Naive Bayes propensity estimator, uniform Naive Bayes propensity estimator, and user propensity estimator respectively. There are two experimental observations. Firstly, the MR method can achieve better performance by engaging several different types of propensity models as shown in the left part of the table. The reason is that different estimations of propensity scores can complement each other and contribute to the multiple robustness of MR beneficially. Secondly, employing the same type of several propensity models, such as two or three NB in the right part of Table 4, tends to harm the model performance. Since the additional same type propensities provide no more useful information but bring redundant noises into the MR learning process.

Sensitivity of Hyper-parameter. By Theorem 5, using a ridge regression form instead of the ordinary least square regression can further control the generalization error bound of MR. We further explore the effect of regularization strength λ in the equation (5) on the debiasing performance of MR. From Figure 1, when varying λ from 0 to 10, the AUC score increases from the beginning, and then declines a little, demonstrating that a proper λ can further improve the performance of MR. Besides, it can effectively mitigate the problem when computing $\hat{\eta}$, especially $\sum_{(u,i)\in\tilde{\mathcal{O}}} \boldsymbol{u}(x_{u,i}; \alpha, \beta) \cdot \boldsymbol{u}^T(x_{u,i}; \alpha, \beta)$, runs into singular matrix.

Conclusion

In this paper, we propose a MR estimator to achieve unbiased learning in RS under weaker conditions. Theoretical analysis demonstrates that the proposed estimator enjoys the property of multiple robustness, i.e., it is unbiased when any of the imputation or propensity models, or a linear combination of these models is accurate. We also show that MR is an enhanced DR estimator when the MR contains only one propensity model and one error imputation model. Further, we analyze the generalization error bound of MR and propose a novel multiple robust learning method with stabilization. We conduct extensive experiments on real-world and semi-synthetic datasets, which shows that the MR estimator can take advantage of multiple candidate propensity models and imputation models to significantly improve performance compared to the state-of-the-art debiasing approaches in RS. In future work, we will study the enhanced versions of MR and focus on sequential training for candidate models in MR.

Acknowledgments

References

Bonner, S.; and Vasile, F. 2018. Causal embeddings for recommendation. In *RecSys*.

Chan, K. C. G. 2013. A simple multiple robust estimator for missing response problem. *Stat*, 2: 143–149.

Chan, K. C. G.; and Yam, S. C. P. 2014. Oracle, Multiple Robust and Multipurpose Calibration in a Missing Response Problem. *Statistical Science*, 29: 380–396.

Chen, J.; Dong, H.; Qiu, Y.; He, X.; Xin, X.; Chen, L.; Lin, G.; and Yang, K. 2021. AutoDebias: Learning to Debias for Recommendation. In *SIGIR*.

Chen, J.; Dong, H.; Wang, X.; Feng, F.; Wang, M.; and He, X. 2020. Bias and Debias in Recommender System: A Survey and Future Directions. *arXiv:2010.03240*.

Chen, S.; and Haziza, D. 2017. Multiply robust imputation procedures for the treatment of item nonresponse in surveys. *Biometrika*, 104: 439–453.

Dai, Q.; Li, H.; Wu, P.; Dong, Z.; Zhou, X.-H.; Zhang, R.; He, X.; Zhang, R.; and Sun, J. 2022. A Generalized Doubly Robust Learning Framework for Debiasing Post-Click Conversion Rate Prediction. In *KDD*.

Ding, S.; Wu, P.; Feng, F.; He, X.; Wang, Y.; Liao, Y.; and Zhang, Y. 2022. Addressing Unmeasured Confounder for Recommendation with Sensitivity Analysis. In *KDD*.

Duan, X.; and Yin, G. 2017. Ensemble Approaches to Estimating the Population Mean with Missing Response. *Scandinavian Journal of Statistics*, 44: 899 –917.

Guo, S.; Zou, L.; Liu, Y.; Ye, W.; Cheng, S.; Wang, S.; Chen, H.; Yin, D.; and Chang, Y. 2021. Enhanced Doubly Robust Learning for Debiasing Post-Click Conversion Rate Estimation. In *SIGIR*.

Han, P. 2016. Combining Inverse Probability Weighting and Multiple Imputation to Improve Robustness of Estimation. *Scandinavian Journal of Statistics*, 43: 246 –260.

Han, P. 2018. Calibration and Multiple Robustness When Data Are Missing Not At Random. *Statistical Science*, 28: 1725–1740.

Han, P.; and Wang, L. 2013. Estimation with missing data: beyond double robustness. *Biometrika*, 100: 417–430.

He, X.; Liao, L.; Zhang, H.; Nie, L.; Hu, X.; and Chua, T. 2017. Neural Collaborative Filtering. In Barrett, R.; Cummings, R.; Agichtein, E.; and Gabrilovich, E., eds., *WWW*.

Jiang, N.; and Li, L. 2016. Doubly Robust Off-policy Value Evaluation for Reinforcement Learning. In Balcan, M.; and Weinberger, K. Q., eds., *ICML*.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In Bengio, Y.; and LeCun, Y., eds., *ICLR*.

Kiyohara, H.; Saito, Y.; Matsuhiro, T.; Narita, Y.; Shimizu, N.; and Yamamoto, Y. 2022. Doubly Robust Off-Policy Evaluation for Ranking Policies under the Cascade Behavior Model. In *WSDM*.

Koren, Y.; Bell, R. M.; and Volinsky, C. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8): 30–37.

Li, H.; Lyu, Y.; Zheng, C.; and Wu, P. 2023a. TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations. In *ICLR*.

Li, H.; Xiao, Y.; Zheng, C.; and Wu, P. 2023b. Balancing Unobserved Confounding with a Few Unbiased Ratings in Debiased Recommendations. In *WWW*.

Li, H.; Zheng, C.; and Wu, P. 2023. StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random. In *ICLR*.

Li, W.; Gu, Y.; and Liu, L. 2020. Demystifying a class of multiple robust estimators. *Biometrika*, 107: 919–933.

Liu, D.; Cheng, P.; Dong, Z.; He, X.; Pan, W.; and Ming, Z. 2020. A general knowledge distillation framework for counterfactual recommendation via uniform data. In *SIGIR*.

Liu, D.; Cheng, P.; Zhu, H.; Dong, Z.; He, X.; Pan, W.; and Ming, Z. 2021. Mitigating Confounding Bias in Recommendation via Information Bottleneck. In *RecSys*.

Marlin, B.; Zemel, R. S.; Roweis, S.; and Slaney, M. 2012. Collaborative filtering and the missing at random assumption. *arXiv:1206.5267*.

Marlin, B. M.; and Zemel, R. S. 2009. Collaborative prediction and ranking with non-random missing data. In *RecSys*.

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Köpf, A.; Yang, E. Z.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; and Chintala, S. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *NeurIPS*.

Saito, Y. 2019. Unbiased Pairwise Learning from Implicit Feedback. In *NeurIPS Workshop*.

Saito, Y. 2020. Doubly robust estimator for ranking metrics with post-click conversions. In *RecSys*.

Saito, Y.; Yaginuma, S.; Nishino, Y.; Sakata, H.; and Nakata, K. 2020. Unbiased Recommender Learning from Missing-Not-At-Random Implicit Feedback. In Caverlee, J.; Hu, X. B.; Lalmas, M.; and Wang, W., eds., *WSDM*.

Sato, M.; Singh, J.; Takemori, S.; Sonoda, T.; Zhang, Q.; and Ohkuma, T. 2019. Uplif-based Evaluation and Optimization of Recommenders. In *RecSys*.

Sato, M.; Takemori, S.; Singh, J.; and Ohkuma, T. 2020. Unbiased Learning for the Causal Effect of Recommendation. In *RecSys*.

Schnabel, T.; Swaminathan, A.; Singh, A.; Chandak, N.; and Joachims, T. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *ICML*.

Steck, H. 2010. Training and testing of recommender systems on data missing not at random. In *KDD*.

Steck, H. 2013. Evaluation of recommendations: rating-prediction and ranking. In *RecSys*.

Swaminathan, A.; and Joachims, T. 2015. The Self-Normalized Estimator for Counterfactual Learning. In Cortes, C.; Lawrence, N. D.; Lee, D. D.; Sugiyama, M.; and Garnett, R., eds., *NeurIPS*.

Tan, Z. 2007. Comment: understanding OR, PS and DR. *Statistical Science*, 22: 560–568.

Wang, X.; Zhang, R.; Sun, Y.; and Qi, J. 2019. Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random. In *ICML*.

Wang, X.; Zhang, R.; Sun, Y.; and Qi, J. 2021. Combating Selection Biases in Recommender Systems with A Few Unbiased Ratings. In *WSDM*.

Wang, Z.; Chen, X.; Wen, R.; Huang, S.-L.; Kuruoglu, E. E.; and Zheng, Y. 2020. Information Theoretic Counterfactual Learning from Missing-Not-At-Random Feedback. *NeurIPS*.

Wu, P.; Han, S.; Tong, X.; and Li, R. 2022a. Propensity score regression for causal inference with treatment heterogeneity. *Statistica Sinica*.

Wu, P.; Li, H.; Deng, Y.; Hu, W.; Dai, Q.; Dong, Z.; Sun, J.; Zhang, R.; and Zhou, X.-H. 2022b. On the Opportunity of Causal Learning in Recommendation Systems: Foundation, Estimation, Prediction and Challenges. In *IJCAI*.

Yang, L.; Cui, Y.; Xuan, Y.; Wang, C.; Belongie, S.; and Estrin, D. 2018. Unbiased offline recommender evaluation for missing-not-at-random implicit feedback. In *RecSys*.

Yuan, B.; Hsia, J.; Yang, M.; Zhu, H.; Chang, C.; Dong, Z.; and Lin, C. 2019. Improving Ad Click Prediction by Considering Non-displayed Events. In *CIKM*.

Zhang, W.; Bao, W.; Liu, X.-Y.; Yang, K.; Lin, Q.; Wen, H.; and Ramezani, R. 2020. Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning. In *WWW*.