Adaptively Estimator Switching for Debiased Recommendation

Jiabo Zou[1]**, Dong Xiao**[1*]
[1] Information Science and Engineering School, Northeastern University
Shenyang 110819, Shenyang, China
mhxyzjb122@foxmail.com, xiaodong@ise.neu.edu.cn[*]
[*] Corresponding Author

Abstract

In the information era, recommendation systems play a crucial role in mitigating information overload by predicting user preferences based on historical interactions. However, traditional recommendation methods often neglect the issue of selection bias arising from non-random missing data, which compromises recommendation quality. To address this, existing approaches such as error imputation-based (EIB), inverse propensity scoring (IPS), and doubly robust (DR) estimators have been proposed. While these methods have demonstrated effectiveness, they suffer from limitations such as sensitivity to small propensity scores, high variance, and inaccuracies in error estimation. This paper introduces a novel switch estimator designed to flexibly integrate the strengths of EIB, IPS, and DR approaches while mitigating their respective weaknesses. Specifically, the proposed method employs a principled Monte Carlo sampling strategy to estimate relative errors in propensity scores and imputation, enabling adaptive threshold-based switching between estimators. This approach ensures robustness to issues arising from small propensity scores and large imputation errors. Experimental evaluations on three real-world datasets demonstrate the superior performance and robustness of the switch estimator in recommendation tasks. The proposed methodology advances the state-of-the-art by offering a practical and effective solution to selection bias in recommendation systems. We conduct experiments on three real-world datasets to show the effectiveness of our method.

Introduction

In the modern information society, the vast amount of information, products, and services enrich users' choices while simultaneously subjecting them to information overload—making it difficult to efficiently access content that best meets their needs. Recommendation systems, as an effective tool to alleviate information overload, are widely applied in areas such as social media, entertainment, short videos, and e-commerce (Alamdari et al. 2020; Liu et al. 2024; Wang et al. 2021a, 2020a; Gong et al. 2022). These systems predict users' preferences for items by combining user characteristics with historical user-item interaction data (Lu et al. 2015). However, data collected through traditional methods often contain selection bias (Chen et al. 2023a). This bias arises because users' feedback on user-item interactions tends to be biased and selective; that is, users are more inclined to provide feedback on content they are interested in while ignoring other content (Pradel, Usunier, and Gallinari 2012). Consequently, recommendation systems often face the problem of non-random missing input data. Traditional recommendation systems typically focus on model structure and the accuracy of predicting observed data. However, ignoring the non-random missingness can significantly enhance the quality of recommendation (Wang et al. 2018; Marlin and Zemel 2009; Steck 2010a).

To mitigate the adverse effects of selection bias, Steck (2010a) proposed an error imputation based (EIB) recommendation method, which first imputes the error of missing data and then uses the observed value and imputed value to train the prediction model. Schnabel et al. (2016a) proposed to use inverse propensity scores to reweight each sample with observed rating. Wang et al. (2019a) proposed to use a doubly robust (DR) method to combine the advantages of EIB and IPS with a joint learning algorithm, which is unbiased when either the imputed errors or propensities are correct for each user-item pair.

However, the EIB method often requires the imputed error to be accurate for obtaining optimal prediction; large biases can lead to significant fluctuations in prediction performance. The IPS algorithm is highly sensitive to small propensity scores, and in real-life scenarios, propensity scores calculated from observed samples often exhibit high variance and extreme value, especially in sparse real-world datasets. Although the DR method combines the advantages of IPS and EIB approaches, it is simultaneously affected by the variance of propensity scores and the bias of estimation errors. This results in large variance and bias when propensity scores are small or estimation errors are inaccurate. To alleviate the issues introduced by the DR estimator, Song et al. (2023) proposed a recommendation method that selectively adjusts estimation bias by setting a threshold to discard the "toxic" imputations by selectively setting the estimation error to zero for samples with large errors of imputation. However, it does not consider the impact of small propensity scores on the model, which is a more practical issue given the sparsity of real-world data.

Building on the aforementioned methods, this paper proposes a switch estimator that can flexibly switch between EIB, IPS, and DR based on the imputed errors and propensity scores. This approach aims to combine the advantages of each method (such as the low variance property of EIB and low bias property of DR) while avoiding potential issues caused by small propensity scores and large errors of imputation. Specifically, inspired by Song et al. (2023), due to we cannot observe the ground truth propensity and prediction error for all user-item pairs, we adopt the principled Monte-Carlo approach to estimate the relative error of estimated value and real value of propensity and imputation and set appropriate thresholds to decide which estimator to use. The main contributions of this paper are summarized below:

- We propose a switch estimator that combines the advantages of EIB, IPS, and DR estimator, which is more robust to smaller propensity and large errors of imputation.
- We adopt a principled Monte-Carlo sampling method for propensity and imputation with a threshold to decide which estimator to switch.
- We conduct experiments on three real-world datasets to show the effectiveness of the proposed switch method.

Related Work

Causal Recommendation

Selection bias is one of the most common biases in the recommender systems (RS) (Luo et al. 2024; Chen et al. 2023b; Wang et al. 2023c; Wu et al. 2022), resulting in the distribution of the observed population being different from the target population. There are many methods proposed to address this issue (Saito and Nomura 2022; Wang et al. 2022b; Zou et al. 2023; Wang et al. 2024, 2023a; Wu et al. 2023). Specifically, methods including the IPS, EIB, and DR were proposed to mitigate the selection bias in RS (Steck 2010b; Saito 2020; Wang et al. 2022a). EIB methods might produce out-of-bound predictions; while the IPS method may suffer from large variance with small propensities (Wang et al. 2022a). DR methods combine the advantages of both the EIB and IPS methods, guaranteeing unbiasedness if either the error imputation model or propensity model is correctly specified. There have been quantities of variants of DR methods to improve the debiasing performance, such as Multi-DR (Zhang et al. 2020), MRDR (Guo et al. 2021), DR-MSE (Dai et al. 2022), BRD-DR (Ding et al. 2022), SDR (Li, Zheng, and Wu 2023), TDR (Li et al. 2023b), DR-V2 (Li et al. 2023d), CDR (Song et al. 2023), KBDR (Li et al. 2024d), N-DR (Li et al. 2024b), DCE-DR (Kweon and Yu 2024), DT-DR (Zhang et al. 2024), UIDR (Li et al. 2024c), and OME-DR (Li et al. 2024d). Besides, recent work on debiased recommendations has been extended to methods using a small amount of unbiased data as a golden standard to correct the misspecified models (Li et al. 2024a, 2023c; Wang et al. 2021b; Chen et al. 2021; Liu et al. 2022). In addition, Wang et al. (Wang et al. 2020b) and Liu et al. (Liu et al. 2023) use information bottleneck-based method and Yang et al. (Yang et al. 2021) and Wang et al. (Wang et al. 2023b) uses adversarial learning for debiasing.

Switch Estimator

Switch estimators are commonly used in fields such as statistical modeling (Fox et al. 2011; Shumway and Stoffer 1991) and reinforcement learning (Thomas and Brunskill 2016; Comanici and Precup 2010). The core idea of switch estimators is to flexibly switch between different estimation methods under varying samples or statistical conditions to balance bias and variance in the estimation process, thereby improving the accuracy and robustness of the estimates. Common types of switch estimators include step-switching estimators, adaptive switching estimators, and hybrid estimators (Thomas and Brunskill 2016). Step-switching estimators switch after a predetermined number of time steps. Farajtabar, Chow, and Ghavamzadeh (2018) proposed an improved doubly robust estimator based on fixed time steps. Adaptive switching estimators dynamically adjust based on certain statistics. Tulsyan et al. (2018) introduced an adaptive switching Bayesian dynamic estimator. An and Yang (2017) proposed a novel state observer with an adaptive switching mechanism by incorporating a switch function matrix into the observer design. Hybrid estimators combine importance sampling and value function estimation, reducing variance through weighted averaging. Building on this, Jiang and Li (2016) proposed a doubly robust estimator that combines importance sampling and value function estimation, effectively reducing estimation variance while maintaining unbiasedness through weighted averaging. Dudık et al. (2014) introduced a hybrid switch estimator that combines importance sampling with regression model weighting, effectively reducing both bias and variance.

Problem Setup

Let $\mathcal{U} = \{u_1, \cdots, u_m\}$ be the users set, $\mathcal{I} = \{i_1, \cdots, i_n\}$ be the item set, and $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ be the set of all user-item pairs. The rating matrix is denoted as $\mathbf{R} \in \mathbb{R}^{m \times n}$ with $r_{u,i}$ as element. Let $o_{u,i} \in \{0, 1\}$ be the observation indicator indicating whether the $r_{u,i}$ is observed and $x_{u,i}$ be the feature. We denote the prediction model as $f_\theta(\cdot)$ parameterized by θ and the predicted ratings as $\hat{r}_{u,i} = f_\theta(x_{u,i})$. The goal is to accurately predict $r_{u,i}$ for all user-item pairs, which can be achieved by minimizing the ideal loss

$$\mathcal{L}_{\text{ideal}}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \mathcal{L}(f_\theta(x_{u,i}), r_{u,i}) := \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e_{u,i},$$

where $\mathcal{L}(\cdot, \cdot)$ is the training loss function such as cross-entropy loss. However, in practice, we cannot obtain the complete rating matrix. We denote the set of user-item pairs with observed ratings as $\mathcal{O} = \{(u, i) \mid o_{u,i} = 1\}$. Thus, the naive method optimizes the average loss over the observed samples

$$\mathcal{L}_{\text{N}}(\theta) = \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} e_{u,i}.$$

Due to the selection bias, $\mathbb{E}\{\mathcal{L}_{\text{N}}(\theta)\} \neq \mathcal{L}_{\text{ideal}}(\theta)$ (Schnabel et al. 2016b; Wang et al. 2019b). Several methods were proposed to unbiasedly estimate the ideal loss, including the EIB, IPS, DR, and their variants. The loss function of EIB method is shown below:

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

where $\hat{e}_{u,i}$ is the error for the imputation model $m(x_{u,i};\phi)$, i.e., $\hat{e}_{u,i} = \mathcal{L}(m(x_{u,i};\phi), \hat{r}_{u,i})$. In addition, the loss function of the IPS method is

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

where $\hat{p}_{u,i}$ is the estimated propensity score for the true exposure probability $p_{u,i} = \Pr(o_{u,i} = 1 \mid x_{u,i})$. The loss function of the vanilla DR method is formulated as

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

When $\hat{e}_{u,i} = 0$, DR degenerates to IPS, and when $\hat{p}_{u,i} = 1$, DR degenerates to EIB. The EIB and IPS estimators are the special case of DR, so we only formulate bias of the DR

$$\mathbf{Bias}\left[\mathcal{L}_{\mathrm{DR}}(\theta)\right] = \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i)\in\mathcal{D}} \frac{(\hat{p}_{u,i} - p_{u,i})}{\hat{p}_{u,i}} (e_{u,i} - \hat{e}_{u,i}) \right|. \tag{1}$$

if either the imputation model or the propensity model is correct, i.e, $\hat{p}_{u,i} = p_{u,i}$ or $\hat{e}_{u,i} = e_{u,i}$, the DR estimator is unbiased. Correspondingly, the variance of DR is

$$\mathbf{Var}\left[\mathcal{L}_{\mathrm{DR}}(\theta)\right] = \frac{1}{|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} \frac{p_{u,i}\left(1 - p_{u,i}\right)}{\hat{p}_{u,i}^2} \left(e_{u,i} - \hat{e}_{u,i}\right)^2. \tag{2}$$

We can easily to obtain $\mathbf{Bias}\left[\mathcal{L}_{\mathrm{EIB}}\right]$ and $\mathbf{Var}\left[\mathcal{L}_{\mathrm{EIB}}\right]$ by setting $\hat{p}_{u,i} = 1$ and $\mathbf{Bias}\left[\mathcal{L}_{\mathrm{IPS}}\right]$ and $\mathbf{Var}\left[\mathcal{L}_{\mathrm{IPS}}\right]$ by setting $\hat{e}_{u,i} = 0$ for all user-item pairs.

Proposed Method

Distinction of Previous Method

Song et al. (2023) finds that if the imputed error $\hat{e}_{u,i}$ extremely deviate from the true prediction error $e_{u,i}$, the $\mathbf{Bias}\left[\mathcal{L}_{\mathrm{IPS}}\right] < \mathbf{Bias}\left[\mathcal{L}_{\mathrm{DR}}\right]$. To be specific, for those user-item pairs with $|\hat{e}_{u,i} - e_{u,i}| > e_{u,i}$, we can reduce the bias by clipping the $\hat{e}_{u,i}$ to zero. The following lemma demonstrates the relations between $e_{u,i}$ and $\hat{e}_{u,i}$

Lemma 1. *Given that $\hat{e}_{u,i}$ and $e_{u,i}$ are independently drawn from two Gaussian distributions $\mathcal{N}\left(\hat{\mu}_{u,i}, \hat{\sigma}_{u,i}^2\right)$ and $\mathcal{N}\left(\mu_{u,i}, \sigma_{u,i}^2\right)$, where $\hat{\mu}_{u,i}$, $\mu_{u,i}, \hat{\sigma}_{u,i}, \sigma_{u,i}$ are bounded with $|\hat{\mu}_{u,i} - \mu_{u,i}| \le \varepsilon_\mu$, $\left|\hat{\sigma}_{u,i}^2 - \sigma_{u,i}^2\right| \le \varepsilon_\sigma^2$, $2\varepsilon_\mu \le \hat{\mu}_{u,i}, m_\mu \le \hat{\mu}_{u,i} \le M_\mu$ and $m_\sigma \le \hat{\sigma}_{u,i} \le M_\sigma$, for any confidence level $\rho (0 \le \rho \le 1)$, the condition $\mathbb{P}\left(|\hat{e}_{u,i} - e_{u,i}| < e_{u,i}\right) \ge \rho$ holds if*

$$\frac{\hat{\sigma}_{u,i}}{\hat{\mu}_{u,i}} < \left(\sqrt{5}\Phi^{-1}(\rho) + \frac{2M_\mu \varepsilon_\sigma}{m_\sigma\left(\sqrt{5}m_\sigma + 2\varepsilon_\sigma\right)} + \frac{2\sqrt{5}\varepsilon_\mu}{\sqrt{5}m_\sigma + 2\varepsilon_\sigma} \right)^{-1}$$

where $\Phi^{-1}(\cdot)$ denotes the inverse of CDF of the standard normal distribution.

Switch Estimator

First, we find that if $\hat{p}_{u,i} \in (0, p_{u,i}/2)$, the absolute value of $(\hat{p}_{u,i} - p_{u,i})/\hat{p}_{u,i}$ is greater than 1. Thus, the bias of the EIB estimator on this user-item pair is less than the DR estimator. Thus, we propose to switch the DR estimator to the EIB estimator in this scenario. Due to we cannot obtain the true propensity $p_{u,i}$ for all user-item pairs. Inspired by Song et al. (2023), we propose the principled Monte Carlo sampling method to control the probability of $\mathbb{P}\left(\hat{p}_{u,i} < p_{u,i}/2\right)$. Specifically, assuming a normal distribution for propensity score is unreasonable because the range of it is between 0 and 1. Thus, we assume the uniform distribution $\mathrm{Unif}(\alpha_{u,i}, 1 - \beta_{u,i})$. Because some users may have overall greater propensities than other users. Thus, we can derive the following theorem for the propensity. In addition, to consider the user heterogeneity, thus we set the propensity threshold at the user level to derive a tighter bound.

Theorem 1. *Given that $\hat{p}_{u,i}$ and $p_{u,i}$ are independently drawn from two uniform distributions $\mathrm{Unif}(\hat{\alpha}_{u,i}, 1 - \hat{\beta}_{u,i})$ and $\mathrm{Unif}(\alpha_{u,i}, 1 - \beta_{u,i})$, where $\hat{\alpha}_{u,i}$, $\alpha_{u,i}, \hat{\beta}_{u,i}, \beta_{u,i}$ are lower bounded with $C_{1,u}$ and upper bounded with $C_{2,u}$, $\alpha_{u,i} + \beta_{u,i} \ge \gamma_u$, $\alpha/2 < 1 - \hat{\beta}$, for any confidence level $\rho (0 < \rho < 1)$, the condition $\mathbb{P}\left(\hat{p}_{u,i} < p_{u,i}/2\right) < \rho$ holds if $\hat{\alpha}_{u,i}/\hat{\beta}_{u,i} < \eta$, where $\eta = 1/C_{2,u} - \frac{4/C_{1,u}}{(1-\gamma_u)\rho} - 1$*

Before the proof, we first introduce the following lemma:

Lemma 2. *Suppose X follows a uniform distribution in $[a, b]$, and Y follows a uniform distribution in $[c, d]$. The difference $Z = X - Y$ has a PDF defined over the interval $[a - d, b - c]$. The PDF is given by*

$$f_Z(z) = \begin{cases} \frac{z-(a-d)}{(b-a)(d-c)}, & \text{for } a - d \le z < c - b \\ \frac{(b-c)-z}{(b-a)(d-c)}, & \text{for } c - b \le z \le b - c \\ 0, & \text{otherwise.} \end{cases}$$

This is a symmetric triangular distribution if $b - a = d - c$, or an asymmetric triangular distribution otherwise.

Proof. Denote $\hat{p}_{u,i} - p_{u,i}/2$ as $z_{u,i}$, from the above lemma, we have the density function of $z_{u,i}$. Then we can derive the probability of $\mathbb{P}\left(z_{u,i} < 0\right)$ based on the bounded parameters, i.e., $|\hat{\alpha}_{u,i} - \alpha_{u,i}/2| \le \varepsilon_{\alpha,u}$, $\left|\hat{\beta}_{u,i} - \beta_{u,i}/2\right| \le \varepsilon_{\beta,u}$.

Specifically, we can calculate the following integration

$$\mathbb{P}\left(z_{u,i} < 0\right) = \int_{\hat{\alpha}-(1-\beta)/2}^{\alpha/2-\left(1-\hat{\beta}\right)} 2 * \frac{u - \hat{\alpha} + (1-\beta)/2}{\left(1 - \hat{\beta} - \hat{\alpha}\right)(1 - \beta - \alpha)} du + \int_{\alpha/2-\left(1-\hat{\beta}\right)}^{0} 2 * \frac{1 - \hat{\beta} - \alpha/2 - u}{\left(1 - \hat{\beta} - \hat{\alpha}\right)(1 - \beta - \alpha)} du$$

We first solve the first part (denoted as I_1):

Denote $u_1^{\text{lower}} = \hat{\alpha} - \frac{1-\beta}{2}$ and $u_1^{\text{upper}} = \frac{\alpha}{2} - (1 - \hat{\beta})$, we get:

$$I_1 = \frac{1}{(1 - \hat{\beta} - \hat{\alpha})(1 - \beta - \alpha)} \left(u_1^{\text{upper}} - u_1^{\text{lower}}\right)^2$$

Table 1: Performance on AUC, NDCG@K, and F1@K on **Coat**, **Yahoo! R3** and **KuaiRec**. The best and the second best results are bolded and underlined, where * means statistically significant results (p-value ≤ 0.05) using the paired-t-test.

	Coat			**Yahoo! R3**			**KuaiRec**		
Method	AUC	NDCG@5	F1@5	AUC	NDCG@5	F1@5	AUC	NDCG@20	F1@20
Naive	$0.703_{\pm0.006}$	$0.605_{\pm0.012}$	$0.467_{\pm0.007}$	$0.673_{\pm0.001}$	$0.635_{\pm0.002}$	$0.306_{\pm0.002}$	$0.753_{\pm0.001}$	$0.449_{\pm0.002}$	$0.124_{\pm0.002}$
IPS	$0.717_{\pm0.007}$	$0.617_{\pm0.009}$	$0.473_{\pm0.008}$	$0.678_{\pm0.001}$	$0.638_{\pm0.002}$	$0.318_{\pm0.002}$	$0.755_{\pm0.004}$	$0.452_{\pm0.010}$	$0.131_{\pm0.004}$
SNIPS	$0.714_{\pm0.012}$	$0.614_{\pm0.012}$	$0.474_{\pm0.009}$	$0.683_{\pm0.002}$	$0.639_{\pm0.002}$	$0.316_{\pm0.002}$	$0.754_{\pm0.003}$	$0.453_{\pm0.004}$	$0.126_{\pm0.003}$
ASIPS	$0.719_{\pm0.009}$	$0.618_{\pm0.012}$	$0.476_{\pm0.009}$	$0.679_{\pm0.003}$	$0.640_{\pm0.003}$	$0.319_{\pm0.003}$	$0.757_{\pm0.005}$	$0.474_{\pm0.007}$	$0.130_{\pm0.005}$
IPS-V2	$0.726_{\pm0.005}$	$0.627_{\pm0.009}$	$0.479_{\pm0.008}$	$0.685_{\pm0.002}$	$0.646_{\pm0.003}$	$0.320_{\pm0.002}$	$0.764_{\pm0.001}$	$0.476_{\pm0.003}$	$0.135_{\pm0.003}$
KBIPS	$0.714_{\pm0.003}$	$0.618_{\pm0.010}$	$0.474_{\pm0.007}$	$0.676_{\pm0.002}$	$0.642_{\pm0.003}$	$0.318_{\pm0.002}$	$0.763_{\pm0.001}$	$0.463_{\pm0.007}$	$0.134_{\pm0.002}$
AKBIPS	$0.732_{\pm0.004}$	$0.636_{\pm0.006}$	$0.483_{\pm0.006}$	$0.689_{\pm0.001}$	$0.658_{\pm0.002}$	$0.324_{\pm0.002}$	$0.766_{\pm0.003}$	$0.478_{\pm0.009}$	$0.138_{\pm0.003}$
DR	$0.718_{\pm0.008}$	$0.623_{\pm0.009}$	$0.474_{\pm0.007}$	$0.684_{\pm0.002}$	$0.658_{\pm0.003}$	$0.326_{\pm0.002}$	$0.755_{\pm0.008}$	$0.462_{\pm0.010}$	$0.135_{\pm0.005}$
DR-JL	$0.723_{\pm0.005}$	$0.629_{\pm0.007}$	$0.479_{\pm0.005}$	$0.685_{\pm0.002}$	$0.653_{\pm0.002}$	$0.324_{\pm0.002}$	$0.766_{\pm0.002}$	$0.467_{\pm0.005}$	$0.136_{\pm0.003}$
MRDR-JL	$0.727_{\pm0.005}$	$0.627_{\pm0.008}$	$0.480_{\pm0.008}$	$0.684_{\pm0.002}$	$0.652_{\pm0.003}$	$0.325_{\pm0.002}$	$0.768_{\pm0.005}$	$0.473_{\pm0.007}$	$0.139_{\pm0.004}$
DR-BIAS	$0.726_{\pm0.004}$	$0.629_{\pm0.009}$	$0.482_{\pm0.007}$	$0.685_{\pm0.002}$	$0.653_{\pm0.002}$	$0.325_{\pm0.003}$	$0.768_{\pm0.003}$	$0.477_{\pm0.006}$	$0.137_{\pm0.004}$
DR-MSE	$0.727_{\pm0.007}$	$0.631_{\pm0.008}$	$0.484_{\pm0.007}$	$0.687_{\pm0.002}$	$0.657_{\pm0.003}$	$0.327_{\pm0.003}$	$0.770_{\pm0.003}$	$0.480_{\pm0.006}$	$0.140_{\pm0.003}$
MR	$0.724_{\pm0.004}$	$0.636_{\pm0.006}$	$0.481_{\pm0.006}$	$0.691_{\pm0.002}$	$0.647_{\pm0.002}$	$0.316_{\pm0.003}$	$0.776_{\pm0.005}$	$0.483_{\pm0.006}$	$0.142_{\pm0.003}$
TDR	$0.714_{\pm0.006}$	$0.634_{\pm0.011}$	$0.483_{\pm0.008}$	$0.688_{\pm0.003}$	$0.662_{\pm0.002}$	$\underline{0.329}_{\pm0.002}$	$0.772_{\pm0.003}$	$0.486_{\pm0.005}$	$0.140_{\pm0.003}$
TDR-JL	$0.731_{\pm0.005}$	$0.639_{\pm0.007}$	$0.484_{\pm0.007}$	$0.689_{\pm0.002}$	$0.656_{\pm0.004}$	$0.327_{\pm0.003}$	$0.772_{\pm0.003}$	$0.489_{\pm0.005}$	$0.142_{\pm0.003}$
StableDR	$0.735_{\pm0.005}$	$0.640_{\pm0.007}$	$0.484_{\pm0.006}$	$0.688_{\pm0.002}$	$0.661_{\pm0.003}$	$0.329_{\pm0.002}$	$0.773_{\pm0.001}$	$0.491_{\pm0.003}$	$0.143_{\pm0.003}$
DR-V2	$0.734_{\pm0.007}$	$0.639_{\pm0.009}$	$0.487_{\pm0.006}$	$0.690_{\pm0.002}$	$0.660_{\pm0.005}$	$0.328_{\pm0.002}$	$0.773_{\pm0.003}$	$0.488_{\pm0.006}$	$0.142_{\pm0.004}$
KBDR	$0.730_{\pm0.003}$	$0.631_{\pm0.005}$	$0.482_{\pm0.006}$	$0.682_{\pm0.002}$	$0.648_{\pm0.003}$	$0.323_{\pm0.002}$	$0.765_{\pm0.004}$	$0.460_{\pm0.006}$	$0.138_{\pm0.003}$
AKBDR	$\mathbf{0.745}_{\pm0.004}$	$0.645_{\pm0.008}$	$0.493_{\pm0.007}$	$\underline{0.692}_{\pm0.002}$	$\underline{0.661}_{\pm0.002}$	$0.328_{\pm0.002}$	$\underline{0.782}_{\pm0.003}$	$\underline{0.498}_{\pm0.008}$	$\mathbf{0.147}_{\pm0.003}$
CDR	$0.743_{\pm0.004}$	$\underline{0.657}_{\pm0.006}$	$\underline{0.495}_{\pm0.005}$	$0.691_{\pm0.002}$	$0.660_{\pm0.002}$	$0.326_{\pm0.003}$	$0.775_{\pm0.004}$	$0.490_{\pm0.009}$	$0.145_{\pm0.003}$
Switch	$\underline{0.741}_{\pm0.002}$	$\mathbf{0.666}^{*}_{\pm0.004}$	$\mathbf{0.501}^{*}_{\pm0.004}$	$\mathbf{0.708}^{*}_{\pm0.003}$	$\mathbf{0.674}^{*}_{\pm0.003}$	$\mathbf{0.337}^{*}_{\pm0.002}$	$\mathbf{0.788}^{*}_{\pm0.003}$	$\mathbf{0.500}_{\pm0.004}$	$\underline{0.146}_{\pm0.003}$

Then we solve the second part (denoted as I_2):

Denote $u_2^{\text{lower}} = \frac{\alpha}{2} - (1-\hat{\beta})$ and $u_2^{\text{upper}} = 0$, we get:

$$I_2 = \frac{3}{(1-\hat{\beta}-\hat{\alpha})(1-\beta-\alpha)}(u_2^{\text{lower}})^2$$

Then we get the final results with $I_1 + I_2$. We know all u_1^{lower}, u_1^{upper}, and u_2^{lower} are less than 1, thus

$$I_1 + I_2 < \frac{4/C_{1,u}}{(1-\gamma_u)\left(1/C_{2,u} - 1 - \hat{\alpha}/\hat{\beta}\right)} < \rho,$$

which means that

$$\hat{\alpha}/\hat{\beta} < 1/C_{2,u} - \frac{4/C_{1,u}}{(1-\gamma_u)\,\rho} - 1$$

□

Then, due to the large user number, instead of pre-specified, we parameterize the η into a neural network $h(x_{u,i})$. In addition, we also parameterize the η in Lemma 1 into a neural network $g(x_{u,i})$. Inspired by Song et al. (2023), we adopt a Monto Carlo sampling strategy. Specifically, we first estimate $\hat{\mu}_{u,i}, \hat{\sigma}_{u,i}, \hat{\alpha}_{u,i}, \hat{\beta}_{u,i}$ for estimating the mean and variance of the imputation and propensity scores by we applying dropout 10 times on the imputation model and propensity model (i.e., randomly omitting 50% of the dimensions of embeddings) and then calculate the mean and variance of $\hat{e}_{u,i}$ and $\hat{p}_{u,i}$ from the dropout model. Then we filter the "toxic" imputation and propensity score based on the condition $\frac{\hat{\sigma}_{ui}}{\hat{\mu}_{ui}} < g(x_{u,i})$ and $\frac{\hat{\alpha}_{u,i}}{\hat{\beta}_{u,i}} < h(x_{u,i})$. We jointly learn the propensity filtering threshold with the propensity model in the propensity learning stage, and learn the imputation filtering threshold, prediction model, and the imputation model in the prediction model learning stage.

Finally, the switch estimator can be formulated as:

$$\mathcal{L}_{\mathbf{Switch}}(\theta) = \begin{cases} \mathcal{L}_{\mathbf{IPS}}(\theta), & \frac{\hat{\sigma}_{ui}}{\hat{\mu}_{ui}} > g(x_{u,i}), \\ \mathcal{L}_{\mathbf{EIB}}(\theta), & \frac{\hat{\alpha}_{u,i}}{\hat{\beta}_{u,i}} > h(x_{u,i}), \\ \mathcal{L}_{\mathbf{DR}}(\theta), & \text{Otherwise.} \end{cases}$$

Experiments

Datasets. We conduct the experiments on three real-world datasets, namely **Coat** (Schnabel et al. 2016b), **Yahoo! R3** (Schnabel et al. 2016b), and **KuaiRec** (Gao et al. 2022), which are widely used in debiased RS because all of them include both biased data and unbiased data. **Coat** dataset consists of 6,960 biased ratings in the training set and 4,640 unbiased ratings in the test set from 290 users and 300 items. The **Yahoo! R3** dataset includes 311,704 biased ratings and 54,000 unbiased from 15,400 users and 1,000 items. Each rating in both datasets are five-scale. We binarize them by letting ratings greater than two to 1 and 0 otherwise. Additionally, we use an industrial dataset **KuaiRec** with 4,676,570 records for video watching ratio of 1,411 users and 3,327 videos. We binarize the records by letting values greater than two be 1 and 0 otherwise.

Baselines. We use matrix factorization as the backbone, and compare our method with the following baselines for comprehensive evaluations: **Naive** method (Marlin and Zemel 2009), IPS-based methods including **IPS** (Schnabel et al. 2016b), **SNIPS** (Schnabel et al. 2016b), **ASIPS** (Saito 2020), **IPS-V2** (Li et al. 2023d), **KBIPS** (Li et al. 2024d) and **AKBIPS** (Li et al. 2024d), and DR-based methods including **DR** (Saito 2020), **DR-JL** (Wang et al. 2019c), **MRDR** (Guo et al. 2021), **DR-BIAS** (Dai et al. 2022), **DR-MSE** (Dai et al. 2022), **MR** (Li et al. 2023a), **TDR** (Li et al. 2023b), **TDR-JL** (Li et al. 2023b), **StableDR** (Li, Zheng, and Wu 2023), **DR-V2** (Li et al. 2023d), **KBDR** (Li et al. 2024d), **AKBDR** (Li et al. 2024d) and **CDR** (Song et al. 2023).

Training Protocols and Details. We tune learning rate in $\{0.01, 0.05\}$ and weight decay in $\{1e-6, 5e-6, 1e-5, \ldots, 1e-3, 5e-3\}$. We use the same hyperparameter search space and follow the results in Li et al. (2024d). In addition, following the previous studies (Guo et al. 2021; Saito 2020; Li et al. 2023d), We evaluate the prediction performance with three widely adopted evaluation metrics: AUC, NDCG@K (N@K), and F1@K, and we set K = 5 on **Coat** and **Yahoo! R3** datasets, and K = 20 on **KuaiRec** dataset.

Experiment Results. Table 1 shows the experiment results on all three datasets. The switch estimator consistently outperformed EIB, IPS, and DR based methods across all datasets. Note that in the Yahoo dataset which has the highest sparsity (2%), the switch estimator demonstrated superior performance, which is due to the robustness of the small propensity and large error of imputations.

Conclusions

In this paper, we have proposed a novel switch estimator for recommendation systems, designed to address the challenges posed by selection bias, small propensity scores, and large errors of imputation. By combining the strengths of existing methods—error imputation-based (EIB), inverse propensity scoring (IPS), and doubly robust (DR)—our approach offers a flexible solution that adapts to the varying conditions of real-world data. The use of a principled Monte Carlo sampling technique allows for the estimation of relative errors in both propensity scores and imputed values, enabling the model to switch between different estimators based on the characteristics of the data. One limitation and future direction is to investigate more flexible switch method, instead of using Monte Carlo Dropout method with a threshold.

Acknowledgement

This work was supported in part by the National Key R&D Program of China under Grant 2022YFB2703304; in part by the National Natural Science Foundation of China under Grant 52074064; in part by the Natural Science Foundation of Science and Technology Department of Liaoning Province under Grant 2021-BS-054; in part by the Fundamental Research Funds for the Central Universities under Grant N2404013, N2404015.

References

Alamdari, P. M.; Navimipour, N. J.; Hosseinzadeh, M.; Safaei, A. A.; and Darwesh, A. 2020. A systematic study on the recommender systems in the E-commerce. *Ieee Access*, 8: 115694–115716.

An, L.; and Yang, G.-H. 2017. Secure state estimation against sparse sensor attacks with adaptive switching mechanism. *IEEE Transactions on Automatic Control*, 63(8): 2596–2603.

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. 2023a. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 41(3): 1–39.

Chen, J.; Dong, H.; Wang, X.; Feng, F.; Wang, M.; and He, X. 2023b. Bias and Debias in Recommender System: A Survey and Future Directions. *ACM Transactions on Information Systems*, 41(3): 1–39.

Comanici, G.; and Precup, D. 2010. Optimal policy switching algorithms for reinforcement learning. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, 709–714.

Dai, Q.; Li, H.; Wu, P.; Dong, Z.; Zhou, X.-H.; Zhang, R.; 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*.

Dudık, M.; Erhan, D.; Langford, J.; and Li, L. 2014. Doubly Robust Policy Evaluation and Optimization. *Statistical Science*.

Farajtabar, M.; Chow, Y.; and Ghavamzadeh, M. 2018. More robust doubly robust off-policy evaluation. In *International Conference on Machine Learning*, 1447–1456. PMLR.

Fox, E.; Sudderth, E. B.; Jordan, M. I.; and Willsky, A. S. 2011. Bayesian nonparametric inference of switching dynamic linear models. *IEEE Transactions on signal processing*, 59(4): 1569–1585.

Gao, C.; Li, S.; Lei, W.; Chen, J.; Li, B.; Jiang, P.; He, X.; Mao, J.; and Chua, T.-S. 2022. KuaiRec: A fully-observed dataset and insights for evaluating recommender systems. In *CIKM*.

Gong, X.; Feng, Q.; Zhang, Y.; Qin, J.; Ding, W.; Li, B.; Jiang, P.; and Gai, K. 2022. Real-time short video recommendation on mobile devices. In *Proceedings of the 31st ACM international conference on information & knowledge management*, 3103–3112.

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*.

Jiang, N.; and Li, L. 2016. Doubly robust off-policy value evaluation for reinforcement learning. In *International conference on machine learning*, 652–661. PMLR.

Kweon, W.; and Yu, H. 2024. Doubly Calibrated Estimator for Recommendation on Data Missing Not At Random. In *WWW*.

Li, H.; Dai, Q.; Li, Y.; Lyu, Y.; Dong, Z.; Zhou, X.-H.; and Wu, P. 2023a. Multiple Robust Learning for Recommendation. In *AAAI*.

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

Li, H.; Wu, K.; Zheng, C.; Xiao, Y.; Wang, H.; Geng, Z.; Feng, F.; He, X.; and Wu, P. 2024a. Removing hidden confounding in recommendation: a unified multi-task learning approach. In *NeurIPS*.

Li, H.; Xiao, Y.; Zheng, C.; and Wu, P. 2023c. Balancing unobserved confounding with a few unbiased ratings in debiased recommendations. In *WWW*, 1305–1313.

Li, H.; Xiao, Y.; Zheng, C.; Wu, P.; and Cui, P. 2023d. Propensity Matters: Measuring and Enhancing Balancing for Recommendation. In *ICML*.

Li, H.; Zheng, C.; Ding, S.; Feng, F.; He, X.; Geng, Z.; and Wu, P. 2024b. Be Aware of the Neighborhood Effect: Modeling Selection Bias under Interference for Recommendation. In *ICLR*.

Li, H.; Zheng, C.; Wang, S.; Wu, K.; Wang, E.; Wu, P.; Geng, Z.; Chen, X.; and Zhou, X.-H. 2024c. Relaxing the Accurate Imputation Assumption in Doubly Robust Learning for Debiased Collaborative Filtering. In *ICML*.

Li, H.; Zheng, C.; Wang, W.; Wang, H.; Feng, F.; and Zhou, X.-H. 2024d. Debiased recommendation with noisy feedback. In *KDD*.

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

Liu, D.; Cheng, P.; Lin, Z.; Luo, J.; Dong, Z.; He, X.; Pan, W.; and Ming, Z. 2022. KDCRec: Knowledge distillation for counterfactual recommendation via uniform data. *IEEE Transactions on Knowledge and Data Engineering*.

Liu, D.; Cheng, P.; Zhu, H.; Dong, Z.; He, X.; Pan, W.; and Ming, Z. 2023. Debiased representation learning in recommendation via information bottleneck. *ACM Transactions on Recommender Systems*, 1(1): 1–27.

Liu, W.; Chen, C.; Liao, X.; Hu, M.; Tan, Y.; Wang, F.; Zheng, X.; and Ong, Y. S. 2024. Learning Accurate and Bidirectional Transformation via Dynamic Embedding Transportation for Cross-Domain Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Lu, J.; Wu, D.; Mao, M.; Wang, W.; and Zhang, G. 2015. Recommender system application developments: a survey. *Decision support systems*, 74: 12–32.

Luo, H.; Zhuang, F.; Xie, R.; Zhu, H.; Wang, D.; An, Z.; and Xu, Y. 2024. A survey on causal inference for recommendation. *The Innovation*.

Marlin, B. M.; and Zemel, R. S. 2009. Collaborative prediction and ranking with non-random missing data. In *Proceedings of the third ACM conference on Recommender systems*, 5–12.

Pradel, B.; Usunier, N.; and Gallinari, P. 2012. Ranking with non-random missing ratings: influence of popularity and positivity on evaluation metrics. In *Proceedings of the sixth ACM conference on Recommender systems*, 147–154.

Saito, Y. 2020. Asymmetric Tri-training for Debiasing Missing-Not-At-Random Explicit Feedback. In *SIGIR*.

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

Saito, Y.; and Nomura, M. 2022. Towards Resolving Propensity Contradiction in Offline Recommender Learning. In *IJCAI*.

Schnabel, T.; Swaminathan, A.; Singh, A.; Chandak, N.; and Joachims, T. 2016a. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*, 1670–1679. PMLR.

Schnabel, T.; Swaminathan, A.; Singh, A.; Chandak, N.; and Joachims, T. 2016b. Recommendations as treatments: Debiasing learning and evaluation. In *ICML*, 1670–1679.

Shumway, R. H.; and Stoffer, D. S. 1991. Dynamic linear models with switching. *Journal of the American Statistical Association*, 86(415): 763–769.

Song, Z.; Chen, J.; Zhou, S.; Shi, Q.; Feng, Y.; Chen, C.; and Wang, C. 2023. CDR: Conservative doubly robust learning for debiased recommendation. In *CIKM*.

Steck, H. 2010a. Training and testing of recommender systems on data missing not at random. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 713–722.

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

Thomas, P.; and Brunskill, E. 2016. Data-efficient off-policy policy evaluation for reinforcement learning. In *International Conference on Machine Learning*, 2139–2148. PMLR.

Tulsyan, A.; Khare, S.; Huang, B.; Gopaluni, B.; and Forbes, F. 2018. A switching strategy for adaptive state estimation. *Signal Processing*, 143: 371–380.

Wang, F.; Zhong, W.; Xu, X.; Rafique, W.; Zhou, Z.; and Qi, L. 2020a. Privacy-aware cold-start recommendation based on collaborative filtering and enhanced trust. In *DSAA*.

Wang, F.; Zhu, H.; Srivastava, G.; Li, S.; Khosravi, M. R.; and Qi, L. 2021a. Robust collaborative filtering recommendation with user-item-trust records. *TCSS*.

Wang, H.; Chang, T.-W.; Liu, T.; Huang, J.; Chen, Z.; Yu, C.; Li, R.; and Chu, W. 2022a. Escm2: Entire space counterfactual multi-task model for post-click conversion rate estimation. In *SIGIR*.

Wang, H.; Kuang, K.; Chi, H.; Yang, L.; Geng, M.; Huang, W.; and Yang, W. 2023a. Treatment effect estimation with adjustment feature selection. In *KDD*.

Wang, H.; Kuang, K.; Lan, L.; Wang, Z.; Huang, W.; Wu, F.; and Yang, W. 2024. Out-of-distribution generalization with causal feature separation. *IEEE Transactions on Knowledge and Data Engineering*, 36(4): 1758–1772.

Wang, H.; Yang, W.; Yang, L.; Wu, A.; Xu, L.; Ren, J.; Wu, F.; and Kuang, K. 2022b. Estimating Individualized Causal Effect with Confounded Instruments. In *KDD*.

Wang, J.; Li, H.; Zhang, C.; Liang, D.; Yu, E.; Ou, W.; and Wang, W. 2023b. Counterclr: Counterfactual contrastive learning with non-random missing data in recommendation. In *ICDM*.

Wang, M.; Gong, M.; Zheng, X.; and Zhang, K. 2018. Modeling dynamic missingness of implicit feedback for recommendation. *Advances in neural information processing systems*, 31.

Wang, W.; Zhang, Y.; Li, H.; Wu, P.; Feng, F.; and He, X. 2023c. Causal Recommendation: Progresses and Future Directions. In *SIGIR*.

Wang, X.; Zhang, R.; Sun, Y.; and Qi, J. 2019a. Doubly robust joint learning for recommendation on data missing not at random. In *International Conference on Machine Learning*, 6638–6647. PMLR.

Wang, X.; Zhang, R.; Sun, Y.; and Qi, J. 2019b. Doubly robust joint learning for recommendation on data missing not at random. In *ICML*.

Wang, X.; Zhang, R.; Sun, Y.; and Qi, J. 2019c. Doubly robust joint learning for recommendation on data missing not at random. In *ICML*.

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

Wang, Z.; Chen, X.; Wen, R.; Huang, S.-L.; Kuruoglu, E.; and Zheng, Y. 2020b. Information theoretic counterfactual learning from missing-not-at-random feedback. In *NeurIPS*, 1854–1864.

Wu, A.; Kuang, K.; Xiong, R.; Li, B.; and Wu, F. 2023. Stable estimation of heterogeneous treatment effects. In *ICML*.

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

Yang, M.; Dai, Q.; Dong, Z.; Chen, X.; He, X.; and Wang, J. 2021. Top-n recommendation with counterfactual user preference simulation. In *CIKM*.

Zhang, H.; Wang, S.; Li, H.; Zheng, C.; Chen, X.; Liu, L.; Luo, S.; and Wu, P. 2024. Uncovering the Propensity Identification Problem in Debiased Recommendations. In *ICDE*.

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*.

Zou, H.; Wang, H.; Xu, R.; Li, B.; Pei, J.; Jian, Y. J.; and Cui, P. 2023. Factual Observation Based Heterogeneity Learning for Counterfactual Prediction. In *CCLR*.