Uncovering the Propensity Identification Problem in Debiased Recommendations

Honglei Zhang
Beijing Jiaotong University
honglei.zhang@bjtu.edu.cn

Shuyi Wang
University of Pennsylvania
sywang25@sas.upenn.edu

Haoxuan Li
Peking University
hxli@stu.pku.edu.cn

Chunyuan Zheng
University of California, San Diego
czheng@ucsd.edu

Xu Chen
Renmin University of China
xu.chen@ruc.edu.cn

Li Liu
Chongqing University
dcsliuli@cqu.edu.cn

Shanshan Luo✉
Beijing Technology and Business University
shanshanluo@btbu.edu.cn

Peng Wu✉
Beijing Technology and Business University
pengwu@btbu.edu.cn

2024 IEEE 40th International Conference on Data Engineering (ICDE) | 979-8-3503-1715-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICDE60146.2024.00056

***Abstract*—In database of recommender systems, users' ratings for most items are usually missing, resulting in selection bias when users selectively choose items to rate. To address this problem, propensity-based methods, e.g., inverse propensity scoring and doubly robust, have been widely studied and applied to missing rating prediction and post-click conversion rate prediction tasks. However, have we completely eliminated the selection bias? Under what missing data mechanism can previous studies completely eliminate the selection bias and lead to unbiased learning? In this paper, following the previous literature on statistics, we first formally define three missing data mechanisms, i.e., missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR), and discuss the widespread prevalence of MNAR in recommender systems. Next, we theoretically reveal that the unbiasedness of previous propensity-based debiasing methods is valid only when data are MCAR or MAR, while it leads to biased predictions when data are MNAR. To tackle this research gap, we propose to disentangle user and item embeddings into the primary latent vector for rating prediction and the auxiliary latent vector for missing mechanism modeling. We prove the identifiablility results, and show that the proposed method can achieve unbiased learning under MNAR with imposed constraints. Extensive experiments are conducted on a semi-synthetic dataset and three real-world datasets, validating the effectiveness of our proposed method.**

***Index Terms*—Recommender system, Identification, Selection bias, Propensity, Missing not at random**

I. Introduction

Recommender systems (RS) aim to solve the problem of information overload by filtering out items that are preferred by users from its log database [1]–[4]. By utilizing structured data stored in the database containing information about user features, item attributes, and interaction records (e.g., clicks and ratings), RS can make personalized recommendations for each user, which has been widely used in areas such as education, business, and entertainment [5]–[8]. However, the interactions recorded in the database were found to suffer from selection bias [9]–[12], due to the (i) *system filtering* strategy at the deployment phase, and (ii) *user selection* preference during the interactive process, which can significantly affect the quality of RS [5], [10], [13], [14]. In addition, selection bias is considered one of the key challenges in many RS tasks, such as rating prediction [6], [15], post-click conversion rate prediction [16]–[18], and post-view click-through&conversion rate prediction [16], [18], [19]. For example, in the rating prediction task, users always choose the favored items to rate, while other unrated items are not considered missing randomly. Hence, there is a significant difference between the distribution of observed events and missing events, causing a severe obstacle for unbiased predictions [15], [20]–[26].

Specifically, the missing mechanisms of selection bias in RS can be divided into three categories [27]–[29]. As shown in Figure 1, the first category is *missing completely at random* (MCAR), that is, the probability of missingness is independent of the observed data (including both the features and the ratings). In such cases, the observed data can be regarded as a representative of the target population, thus the selection bias vanishes. The second category is *missing at random* (MAR), in which the probability of missingness depends only on the features of the users and items, which can be fully observed. Most studies that eliminate selection bias are limited to data MAR [30]–[32]. The third category is *missing not at random* (MNAR), that is, the probability of missingness depends on both the (observed) features and the (partially observed) ratings, where missing ratings pose a serious challenge for current debiasing methods.

Traditional methods for addressing selection bias mainly fall into two groups. The first very diverse group includes a data imputation module which aims at estimating the missing data. These techniques employ heuristic methods [33] as well as learning-based models [15]. The second group utilizes inverse propensity score (IPS) methods to reweight observed data during training, based on the estimated probability of their observation [6]. To capitalize on the advantages offered by both approaches, the doubly robust (DR) joint learning [15]

2375-026X/24/$31.00 ©2024 IEEE
DOI 10.1109/ICDE60146.2024.00056

is proposed to effectively reduce the estimation error. Despite the efforts of existing propensity-based methods to mitigate selection bias, ratings are almost MNAR, not MAR, and also the learned propensities are easily mis-specified [34], which leads to the lack of identifiability and unbiasedness guarantees.

In light of this, we raise several research questions which aim at uncovering the limitations of eliminating selection bias in RS. Could the existing propensity-based debiasing methods ensure the identifiability of the propensity when data is MNAR? Besides, under what missing data mechanism can previous studies completely eliminate the selection bias and lead to unbiased learning?

To answer the above questions, in this paper, we first reveal that the unbiasedness of previous propensity-based debiasing methods holds only when data are MCAR or MAR, while it leads to biased predictions when data are MNAR. We then provide a comprehensive discussion on how to determine the parameters of interest when learning propensities to enable unbiased learning under all three missing data mechanisms. Based on the parameters of interest under data MNAR, we propose a novel disentanglement approach based on collaborative filtering, which ensures the identifiability in MNAR scenarios. Such identifiability guarantees the uniqueness and the learnability of the propensity model when the auxiliary variables are correctly disentangled, then we have sufficient information for accurate estimation of the propensities.

We further propose a multi-task learning method to disentangle auxiliary embeddings and learn propensities simultaneously under data MNAR. In contrast to previous studies with hard sharing or no sharing of embeddings between the propensity and the prediction model, the proposed method *partially* shares the embeddings of the propensity model and the prediction model to estimate propensities. Meanwhile, the training signals of the unsharing part of the propensity model are given by the disentangled embeddings. We also theoretically demonstrate that the proposed method inherits identifiability and is unbiased under data MNAR. Experiments on a semi-synthetic dataset and three real-world datasets validate the effectiveness of the proposed method.

The contributions of this paper can be summarized below:

- We formally define the three missing da ta mechanisms and show that the previous propensity-based debiasing methods are unbiased only under data MAR or MCAR.
- We develop a novel disentanglement approach to the propensity model by introducing an auxiliary variable, which ensures the identifiability of the learned propensities and unbiasedness under data MNAR.
- We further propose a multi-task learning method, which partially shares the embeddings between the prediction model and the propensity model to learn the auxiliary embeddings and the MNAR propensities simultaneously.
- We conduct extensive experiments on one semi-synthetic dataset and three real-world datasets, including a publicly available large-scale industrial dataset, and the experimental results demonstrate the superiority of the proposed method in recommendation accuracy and efficiency.

II. Problem Setting and Notation

We now formulate the rating prediction task in the presence of selection bias. Let $\mathcal{U} = \{u_1, u_2, ..., u_M\}$ and $\mathcal{I} = \{i_1, i_2, ..., i_N\}$ be the sets of M users and N items, respectively. For each user-item pair (u, i), let $x_{u,i}$ be the feature vector of user u and item i, and $r_{u,i}$ be the rating of user u on item i. In fact, users select only a small subset of items to rate, and most of the ratings are missing. Define $o_{u,i}$ as the observing indicator of $r_{u,i}$, where $o_{u,i} = 1$ means that the rating $r_{u,i}$ is observed in the collected data and $o_{u,i} = 0$ otherwise. Let $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ be the set of all user-item pairs and $\mathcal{O} = \{(u, i) \in \mathcal{D} : o_{u,i} = 1\}$ be the set of observed user-item pairs, we aim to predict the ratings $r_{u,i}$ for all $(u, i) \in \mathcal{D}$.

Let $\hat{r}_{u,i} = f(x_{u,i}; \theta)$ be a rating prediction model parameterized by θ. Ideally, if all ratings were observed, the ideal loss function for training the prediction model is defined as

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

where $e_{u,i}$ is the prediction error. For example, the squared loss $e_{u,i} = (\hat{r}_{u,i} - r_{u,i})^2$. However, optimizing $\mathcal{L}_{ideal}(\theta)$ is infeasible due to the missing ratings $\{r_{u,i} : o_{u,i} = 0\}$.

A naive method for training the prediction model is to discard the missing events directly and only use the observed ratings. The corresponding loss is

$$\mathcal{E}_{Naive}(\theta) = \frac{1}{|\mathcal{O}|} \sum_{(u,i)\in\mathcal{O}} e_{u,i}. \tag{2}$$

It is well known that $\mathcal{E}_{Naive}$ is an unbiased estimator of $\mathcal{L}_{ideal}(\theta)$ only when the missing mechanism is MCAR, that is, the observed user-item pairs can represent the target population. However, due to the presence of selection bias, the data collected in RS cannot be used as a representative sample of the entire user-item matrix, which is regarded as our target population, resulting in $\mathcal{E}_{Naive}(\theta)$ being biased.

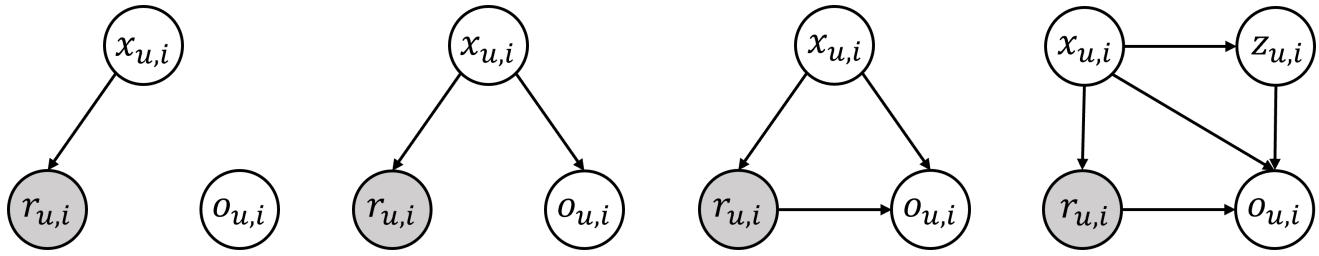
To tackle this problem, many debiasing methods have been proposed [9], [35], but notably, the unbiasedness of these methods is highly dependent on the missing data mechanisms. In the following, we describe the various missing mechanisms via causal graphs and then the debiasing methods in details.

III. Analysis of missing data mechanism

In this section, we first introduce the definition of causal graphs [36], [37], which helps to understand and formally analyze three missing data mechanisms in RS [27]–[29], [38]. Then we show that the existing propensity-based debiasing methods are biased under data MNAR.

A. Missing Mechanisms Illustration via Causal Graphs

Causality aims to study causal dependence rather than correlation between variables, which has gained considerable attention in the database community in recent years [39], [40], such as recommender systems [1], [41] and SQL queries [42], [43]. Causal graph offers an effective tool to intuitively describe causal relationships between variables by using a



(a) MCAR (b) MAR (c) MNAR (d) Disentangling on MNAR

Fig. 1: Typical causal graphs for different missing mechanisms, where the grey circle indicates that the variable is partially observed. The proposed method also works if the arrow direction between $x_{u,i}$ and $z_{u,i}$ is reversed.

directed acyclic graph, where nodes represent variables and directed edges represent potential direct causal effects [36], [37]. When there are no edges, it means that there is no direct effect between the variables.

Figures 1(a)-(c) show the causal graphs of three missing data mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR), under which the challenge of achieving unbiased learning increases. The causal graph consists of three variables: the user and item features $x_{u,i}$, the rating of user to the item $r_{u,i}$, and the rating observing indicator $o_{u,i}$. In RS, users usually have different preferences for items, leading to different ratings, which can be expressed as $x_{u,i} \rightarrow r_{u,i}$. To collect MCAR data, we may ask users to rate randomly selected items, thus neither $x_{u,i}$ nor $r_{u,i}$ have a causal effect on $o_{u,i}$. Compared to the MCAR, the MAR considers the influence of user and item features on the rating indicator, accounting for the self-selection by the users, denoted as $x_{u,i} \rightarrow o_{u,i}$. As the most general missingness mechanism, MNAR further considers the effect of user preferences (reflected in ratings) on sample selection, e.g., some users choose their favorite items to rate, which can be given as $r_{u,i} \rightarrow o_{u,i}$.

B. Missing Mechanisms Analysis

MCAR: the observing indicator $o_{u,i}$ is completely independent of the rating $r_{u,i}$ and feature $x_{u,i}$, i.e., $\mathbb{P}(o_{u,i} = 1|x_{u,i}, r_{u,i}) = \mathbb{P}(o_{u,i} = 1)$ is a constant, or equivalently, $o_{u,i} \perp\!\!\!\perp (r_{u,i}, x_{u,i})$. In RS, the data collected by a random logging policy or A/B test [44]–[47] is MCAR, since the observing indicator can be regarded as the results of a random coin toss, and thus unrelated to both the rating and feature. For MCAR data, the distributions of observed events and missing events are the same, thus the selection bias vanishes and the naive estimator $\mathcal{E}_{Naive}(\theta)$ is unbiased.

MAR: the missing probability depends only on the fully observed features $x_{u,i}$, i.e., $\mathbb{P}(o_{u,i} = 1|x_{u,i}, r_{u,i}) = \mathbb{P}(o_{u,i} = 1|x_{u,i})$, or equivalently, $o_{u,i} \perp\!\!\!\perp r_{u,i} \mid x_{u,i}$. In RS, the data collected by a given recommendation policy are MAR, because the recommendation policy decides whether to expose an item to a user relies only on the fully observed features. A typical task is post-view click-through rate (pCTR) prediction [48], where whether an item is exposed to a user depends on a given exposure policy based on the information observed about the item and the user.

MNAR: the missing probability depends both the features $x_{u,i}$ and the unobserved rating $r_{u,i}$, i.e., $\mathbb{P}(o_{u,i} = 1|x_{u,i}, r_{u,i}) \neq \mathbb{P}(o_{u,i} = 1|x_{u,i})$, or equivalently, $o_{u,i} \not\perp\!\!\!\perp r_{u,i} \mid x_{u,i}$. In RS, data MNAR is the most common case, as users are free to choose which items to interact with. For example, in the task of rating prediction [6], [49], [50], users prefer to rate items they like, and thus, whether a user's rating is observed depends on the value of the rating itself, thus the collected ratings are MNAR.

In a nutshell, MCAR is the simplest case where there is no bias, and the naive estimator is sufficient to achieve unbiasedness. For MAR, the missing probability is only based on fully observed features $x_{u,i}$, thus it can be accurately estimated by modeling $o_{u,i}$ using $x_{u,i}$. Unfortunately, it is very difficult to achieve unbiasedness under data MNAR, since many ratings $r_{u,i}$ are missing and have a direct effect on $o_{u,i}$.

One may argue that we can first impute the missing ratings based on the observed ratings and then model the observing probability $\mathbb{P}(o_{u,i} = 1|x_{u,i}, r_{u,i})$ using both the observed ratings and imputed ratings. However, such a strategy also fails under data MNAR. Specifically, the missing data mechanism $o_{u,i} \not\perp\!\!\!\perp r_{u,i} \mid x_{u,i}$ implies that

$$\mathbb{P}(r_{u,i}|x_{u,i}, o_{u,i} = 1) \neq \mathbb{P}(r_{u,i}|x_{u,i}, o_{u,i} = 0),$$

which indicates the imputation model trained based on the observed events $\mathcal{O}$ cannot extrapolate to the missing events $\mathcal{D} \setminus \mathcal{O}$, i.e., imputing the missing ratings is infeasible. This further demonstrates the difficulty of achieving unbiased learning when using propensity-based methods under data MNAR.

C. Previous Propensity-Based Estimators under Data MNAR

Many propensity-based methods have tried to develop unbiased estimators of the ideal loss $\mathcal{L}_{ideal}(\theta)$, and then train the prediction model by minimizing the estimated loss. For example, the inverse propensity score (IPS) estimator is

$$\mathcal{E}_{IPS}(\theta) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}}, \tag{3}$$

TABLE I: Unbiasedness of MCAR, MAR, and MNAR propensities under different missing data mechanisms, where ✓ and × denote the unbiased and biased results, respectively.

Propensity	MCAR	MAR	MNAR
MCAR propensity $\mathbb{P}(o=1)$	✓	×	×
MAR propensity $\mathbb{P}(o=1\|x)$	✓	✓	×
MNAR propensity $\mathbb{P}(o=1\|x,r)$	✓	✓	✓

where $\hat{p}_{u,i}$ is an estimate of propensity score $p_{u,i} \triangleq \mathbb{P}(o_{u,i}=1|x_{u,i})$. By further imputing the prediction errors, the doubly robust (DR) estimator is

$$\mathcal{E}_{DR}(\theta) = \frac{1}{|\mathcal{D}|}\sum_{(u,i)\in\mathcal{D}}\Big[\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i}-\hat{e}_{u,i})}{\hat{p}_{u,i}}\Big], \quad (4)$$

where $\hat{e}_{u,i}$ is an estimate of $g_{u,i} \triangleq \mathbb{E}[e_{u,i}|x_{u,i}]$, i.e., it predicts $e_{u,i}$ using $x_{u,i}$. IPS and DR methods show competing performance in practice and have been widely studied for debiasing [6], [15], [17], [35], [51], [52], with the following conditions for achieving unbiasedness under data MAR.

Lemma 1. *Under data MAR, $\mathcal{E}_{IPS}(\theta)$ is an unbiased estimator of $\mathcal{L}_{ideal}(\theta)$ if all learned propensities are accurate, i.e., $\hat{p}_{u,i}=p_{u,i}$, and $\mathcal{E}_{DR}(\theta)$ is an unbiased estimator if either all the learned propensities or all the imputed errors are accurate, i.e., $\hat{p}_{u,i}=p_{u,i}$ or $\hat{e}_{u,i}=g_{u,i}$.*

However, $\mathcal{E}_{IPS}(\theta)$ and $\mathcal{E}_{DR}(\theta)$ fail to achieve unbiasedness under data MNAR, because the underlying target propensity is mismatched to the missing data mechanism. Table I summarizes the correspondence between target propensity, missing mechanisms, and unbiasedness of propensity-based estimators. Specifically, we define the MNAR propensity $\tilde{p}_{u,i}$ as

$$\tilde{p}_{u,i} \triangleq \mathbb{P}(o_{u,i}=1|x_{u,i},r_{u,i}), \quad (5)$$

while call $p_{u,i}=\mathbb{P}(o_{u,i}=1|x_{u,i})$ the MAR propensity. Intuitively, the observing indicator $o_{u,i}$ depends on both $x_{u,i}$ and $r_{u,i}$ under data MNAR, while the MAR propensity $p_{u,i}$ is a function of only $x_{u,i}$ that is not able to capture the influence of $r_{u,i}$ on $o_{u,i}$. Therefore, the MAR propensity $p_{u,i}$ leads to biased predictions when using the previous IPS and DR estimators under data MNAR. However, the unbiasedness of the IPS and DR estimators under data MNAR can be achieved if the MNAR propensity is used.

Lemma 2. *(a) Under data MNAR, both the IPS and DR estimators are biased when using the MAR propensity $\hat{p}_{u,i}=p_{u,i}$; (b) Under data MNAR, both the IPS and DR estimators are unbiased when using the MNAR propensity $\hat{p}_{u,i}=\tilde{p}_{u,i}$.*

In practice, we usually train the propensity model by fitting $o_{u,i}$ using the fully observed $x_{u,i}$, such as logistic regression or a neural network [6], [50], [52] with variance regularization [17], [51], [53]. However, previous studies only focus on estimating the MAR propensity instead of the MNAR propensity, which leads to biased estimations. In fact, obtaining accurate estimates of the MNAR propensity by fitting $o_{u,i}$ with $x_{u,i}$ is a very difficult task due to the partially missing $r_{u,i}$, which can be illustrated from a geometric perspective: $\tilde{p}_{u,i}$ is a point on the space spanned by both $x_{u,i}$ and $r_{u,i}$, while $\hat{p}_{u,i}$ (obtained by modeling $o_{u,i}$ with $x_{u,i}$) is a point on the space spanned by $x_{u,i}$. Therefore, $\hat{p}_{u,i}$ always has a gap with $\tilde{p}_{u,i}$ as long as $r_{u,i}$ has a non-zero effect on $o_{u,i}$.

IV. Proposed Method

In this section, we first reveal the identifiability problem under data MNAR, and then propose a novel method to disentangle an auxiliary variable to address the identifiability problem and obtain unbiased learning. Finally, we further propose a multi-task learning approach to adaptively disentangle a learnable auxiliary representation.

A. Identifiability Problem under Data MNAR

The essential reason for the failure of previous IPS and DR estimators under data MNAR is the existence of the identifiability problem. An estimand satisfies identifiability if it can be written as a function of the distribution of the observed data, which implies that one can use observed data to obtain an (asymptotically) unbiased estimation [29].

Formally, the aim of debiased recommendation is to train a prediction model $\hat{r}_{u,i}=f(x_{u,i};\theta)$ to predict $r_{u,i}$, which is equivalent to estimating $\mathbb{P}(r_{u,i}|x_{u,i})$, as the estimand of interest. Given the pre-defined estimand, our goal is to learn it based on the observed data $\mathbb{P}(o_{u,i}=1,r_{u,i}|x_{u,i})$, which can be decomposed as

$$\mathbb{P}(o_{u,i}=1,r_{u,i}|x_{u,i}) = \mathbb{P}(o_{u,i}=1|x_{u,i},r_{u,i})\mathbb{P}(r_{u,i}|x_{u,i}).$$

This equation indicates that the MNAR propensity $\mathbb{P}(o_{u,i}=1|x_{u,i},r_{u,i})$ could bridge the gap between the observed data $\mathbb{P}(o_{u,i}=1,r_{u,i}|x_{u,i})$ and the target estimand $\mathbb{P}(r_{u,i}|x_{u,i})$.

Under data MAR, i.e., $r_{u,i} \perp\!\!\!\perp o_{u,i}|x_{u,i}$, we have $\mathbb{P}(o_{u,i}=1|x_{u,i},r_{u,i}) = \mathbb{P}(o_{u,i}=1|x_{u,i})$. Note that both $o_{u,i}$ and $x_{u,i}$ are fully observed, thus $\mathbb{P}(o_{u,i}=1|x_{u,i})$ can be derived from the observed data, so does the estimand $\mathbb{P}(r_{u,i}|x_{u,i})$. However, under data MNAR, estimating $\mathbb{P}(r_{u,i}|x_{u,i})$ suffers from the identifiability problem and ignoring it will lead to biased estimates, as shown in the following example.

Example 1. *Consider the following two sets of models:*

$$\textit{model (a)}\begin{cases}\mathbb{P}_1(o_{u,i}=1|r_{u,i},x_{u,i})=\sigma(-4+2r_{u,i})\\ \mathbb{P}_1(r_{u,i}|x_{u,i})=\phi(r_{u,i}-1),\end{cases}$$

$$\textit{model (b)}\begin{cases}\mathbb{P}_2(o_{u,i}=1|r_{u,i},x_{u,i})=\sigma(4-2r_{u,i})\\ \mathbb{P}_2(r_{u,i}|x_{u,i})=\phi(r_{u,i}-3),\end{cases}$$

where $\sigma(\cdot)$ is the sigmoid function, $\phi(\cdot)$ is the density of the standard normal distribution.

We show that one *cannot* distinguish whether the true MNAR propensity is $\mathbb{P}_1(o_{u,i}=1|r_{u,i},x_{u,i})$ or $\mathbb{P}_2(o_{u,i}=1|r_{u,i},x_{u,i})$ based on the observed data, i.e., $\tilde{p}_{u,i}$ suffers from the identifiability problem.

The reason is that models (a) and (b) can generate the same observed data distribution that

$$\begin{aligned}&\mathbb{P}_1(o_{u,i}=1|r_{u,i},x_{u,i})\mathbb{P}_1(r_{u,i}|x_{u,i})\\&=\mathbb{P}_2(o_{u,i}=1|r_{u,i},x_{u,i})\mathbb{P}_2(r_{u,i}|x_{u,i}),\end{aligned}\tag{6}$$

which implies that both model (a) and model (b) are the solution that maximizes the likelihood function of the observed data, since they produce the same distribution of observed data. Thus, the solution obtained by optimizing likelihood function has a positive probability of being an incorrect model, which leads to biased estimates of the ideal loss. We provide the detailed analysis as shown below: the equation in Example 1

$$\begin{aligned}&\mathbb{P}_1(o_{u,i}=1|r_{u,i},x_{u,i})\mathbb{P}_1(r_{u,i}|x_{u,i})\\&=\mathbb{P}_2(o_{u,i}=1|r_{u,i},x_{u,i})\mathbb{P}_2(r_{u,i}|x_{u,i})\end{aligned}$$

is equivalent to

$$\begin{aligned}&\frac{\phi(r_{u,i}-3)\exp(-2r_{u,i}+4)}{1+\exp(-2r_{u,i}+4)}=\frac{\phi(r_{u,i}-1)\exp(2r_{u,i}-4)}{1+\exp(2r_{u,i}-4)}\\&\Leftrightarrow\exp\left\{\frac{(r_{u,i}-1)^2-(r_{u,i}-3)^2}{2}\right\}=\frac{1+\exp(2r_{u,i}-4)}{1+\exp(-2r_{u,i}+4)}\\&\Leftrightarrow\exp(2r_{u,i}-4)\left\{1+\exp(-2r_{u,i}+4)\right\}=1+\exp(2r_{u,i}-4)\\&\Leftrightarrow\exp(2r_{u,i}-4)+1=1+\exp(2r_{u,i}-4).\end{aligned}$$

In summary, identifiability is a prerequisite for correct and accurate estimation of $\tilde{p}_{u,i}$. If it is not satisfied, we can never know if we are learning a correct model. It further indicates that accurately estimating $\tilde{p}_{u,i}$ under MNAR is a difficult task without additional information [54], [55].

Fortunately, as shown in Figure 1(d), we show the identifiability of $\tilde{p}_{u,i}$ can be recovered, if we have an extra auxiliary variable $z_{u,i}$ that satisfies the conditions in Assumption 1, as shown in the following Lemma 3.

Assumption 1. *(i)* $z_{u,i} \perp\!\!\!\perp r_{u,i} \mid x_{u,i}$*; (ii)* $z_{u,i} \not\perp\!\!\!\perp o_{u,i} \mid x_{u,i}$.

Lemma 3. *Under data MNAR, if there exists an auxiliary variable* $z_{u,i}$ *satisfies Assumption 1, then the joint distribution of* $(x_{u,i}, r_{u,i}, z_{u,i}, o_{u,i})$ *is identifiable if and only if*

$$\frac{\mathbb{P}_1(o_{u,i}=1 \mid z_{u,i},x_{u,i},r_{u,i})}{\mathbb{P}_2(o_{u,i}=1 \mid z_{u,i},x_{u,i},r_{u,i})}\neq\frac{\mathbb{P}_2(r_{u,i} \mid x_{u,i})}{\mathbb{P}_1(r_{u,i} \mid x_{u,i})},\tag{7}$$

where $\mathbb{P}_j(o_{u,i}|z_{u,i},x_{u,i},r_{u,i})$ *and* $\mathbb{P}_j(r_{u,i}|x_{u,i})$ *for* $j=1,2$ *are any two distinct candidate distributions.*

Lemma 3 presents a necessary and sufficient condition for identifiability of the joint distribution of the full data. That is, the observed distribution $\mathbb{P}(o_{u,i}=1, r_{u,i}, z_{u,i}, x_{u,i})$ and the joint distribution $\mathbb{P}(o_{u,i}, r_{u,i}, z_{u,i}, x_{u,i})$ are one-to-one correspondence, which avoids the unidentifiable scenario in Example 1 and ensures that it is possible to recover the joint distribution from the observed distribution. Based on this, the MNAR propensity can be uniquely determined by the observed data, which further provides the theoretic basis of the unbiased estimation for our proposed method.

In addition, condition (7) in Lemma 3 is a relatively mild condition, and many commonly used parametric and semi-parametric models satisfy this condition, such as the separable logistic and Probit models. In general, when the ratio $\mathbb{P}_1(o_{u,i} \mid z_{u,i},x_{u,i},r_{u,i})/\mathbb{P}_2(o_{u,i} \mid z_{u,i},x_{u,i},r_{u,i})$ is either a constant or varies with $z_{u,i}$ for any two distinct candidate distributions, the condition (7) always holds, which provides a guidance to guarantee the identifiability of joint distribution. Below, we clarify that the separable logistic model satisfies the condition (7). For simplicity, we omit covariates $x_{u,i}$ in the following theorem.

Theorem 1 (Identifiability under MNAR)**.** *Consider the separable logistic missing data mechanism:*

$$\mathbb{P}(o_{u,i}=1 \mid z_{u,i},r_{u,i})=\frac{\exp\{q(z_{u,i})+g(r_{u,i})\}}{1+\exp\{q(z_{u,i})+g(r_{u,i})\}}.\tag{8}$$

where $q(\cdot)$ *and* $g(\cdot)$ *are unknown differentiable functions. For example,* $q(\cdot)$ *and* $g(\cdot)$ *can take the parametric forms:*

$$q(z_{u,i})=\alpha_0+\alpha_1 z_{u,i},\ g(r_{u,i})=\beta_0+\beta_1 r_{u,i},$$

where α_j *and* β_j *are some unknown parameters. The above model is separable since the conditional probability* $\mathbb{P}(o_{u,i}=1 \mid z_{u,i},r_{u,i})$ *excludes an interaction between* $r_{u,i}$ *and* $z_{u,i}$*, making equation* (8) *satisfies the condition* (7)*.*

Proof Sketch. Suppose there exist two density functions that make the ratios equal, that is,

$$\frac{\operatorname{expit}\left\{q_1(z_{u,i})+g_1(r_{u,i})\right\}}{\operatorname{expit}\left\{q_2(z_{u,i})+g_2(r_{u,i})\right\}}=h(r_{u,i})\tag{9}$$

for some function $h(r_{u,i})$, which is a function depends on $r_{u,i}$ only. Taking the derivative of $z_{u,i}$ on both sides, and we have

$$\begin{aligned}&\frac{\partial q_1(z_{u,i})/\partial z_{u,i}}{\partial q_2(z_{u,i})/\partial z_{u,i}}\left[1+\exp\left\{q_2(z_{u,i})+g_2(r_{u,i})\right\}\right]\\&=1+\exp\left\{q_1(z_{u,i})+g_1(r_{u,i})\right\}.\end{aligned}\tag{10}$$

Taking the derivative of $r_{u,i}$ on both sides, then we have

$$\begin{aligned}&\frac{\partial q_1(z_{u,i})/\partial z_{u,i}}{\partial q_2(z_{u,i})/\partial z_{u,i}}\exp\left\{q_2(z_{u,i})-q_1(z_{u,i})\right\}\\&=\frac{\partial g_1(r_{u,i})/\partial r_{u,i}}{\partial g_2(r_{u,i})/\partial r_{u,i}}\exp\left\{g_1(r_{u,i})-g_2(r_{u,i})\right\}.\end{aligned}$$

Note that the left-hand side of the above equation is a function of $z_{u,i}$, but the right side is a function of $r_{u,i}$, therefore, we must have

$$\frac{\partial q_1(z_{u,i})/\partial z_{u,i}}{\partial q_2(z_{u,i})/\partial z_{u,i}}\exp\left\{q_2(z_{u,i})-q_1(z_{u,i})\right\}=c_1,$$

for some constant c_1. Multiply both sides of equation (10) by $\exp\{-q_1(z_{u,i})\}$, and then

$$\begin{aligned}&c_1\left[\exp\left\{-q_2(z_{u,i})\right\}+\exp\left\{g_2(r_{u,i})\right\}\right]\\&=\exp\left\{-q_1(z_{u,i})\right\}+\exp\left\{g_1(r_{u,i})\right\},\end{aligned}$$

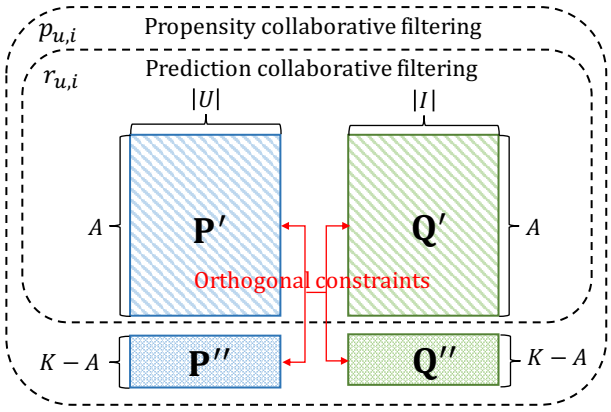


Fig. 2: The proposed disentanglement approach.

and there is some constant c_2 that satisfies

$$c_1 \exp\{-q_2(z_{u,i})\} + c_2 = \exp\{-q_1(z_{u,i})\},$$
$$c_1 \exp\{g_2(r_{u,i})\} - c_2 = \exp\{g_1(r_{u,i})\}.$$

By substituting $q_2(z_{u,i})$and $g_2(r_{u,i})$ in equation (9) with the above expression, we get

$$h(r_{u,i}) = 1 + c_2 \exp\{-g_1(r_{u,i})\}.$$

Note that $1 + c_2 \exp\{-g_1(r_{u,i})\} > 1$ for $c_2 > 0$, and $1 + c_2 \exp\{-g_1(r_{u,i})\} < 1$ for $c_2 < 0$. Besides, note that $h(r_{u,i})$ is the ratio of two densities, i.e. $h(r_{u,i}) = f_2(r_{u,i})/f_1(r_{u,i})$, and so we must have $c_2 = 0$, and thus $h(r_{u,i}) = 1$. As a result, the joint distribution is identified. □

Theorem 1 provides a concrete case that is identifiable under data MNAR. Therefore, to achieve unbiased learning under data MNAR, it is sufficient to find an auxiliary variable $z_{u,i}$ that satisfies Assumption 1 and condition in equation (7).

Notably, different from the condition in equation (6), *the condition in equation (7) is much weaker and many commonly used models satisfy this condition,* such as Logistic regression and Poisson factorization [56]. In general, condition in equation (7) can be verified based on the adopted model and typically holds when $z_{u,i}$ and $r_{u,i}$ have no interaction effects on $o_{u,i}$. More specific, the condition holds when the ratio $\mathbb{P}_1(o_{u,i} \mid z_{u,i}, x_{u,i}, r_{u,i})/\mathbb{P}_2(o_{u,i} \mid z_{u,i}, x_{u,i}, r_{u,i})$ is either a constant or varies with $z_{u,i}$ for any two distinct distributions.

B. Disentangle the Auxiliary Variable

To address the identifiability problem of MNAR propensity, we propose a novel method that disentangles an auxiliary variable $z_{u,i}$ that satisfies Assumption 1 based on collaborative filtering. Specifically, as shown in Figure 2, we disentangle the feature embedding $[\mathbf{p_u}, \mathbf{q_i}]$ into two parts $x_{u,i}$ and $z_{u,i}$, where $\mathbf{p_u} = [\mathbf{p'_u}, \mathbf{p''_u}]$ is the feature embedding for user u, $\mathbf{q_i} = [\mathbf{q'_i}, \mathbf{q''_i}]$ is the feature embedding for item i, $x_{u,i} = [\mathbf{p'_u}, \mathbf{q'_i}]$ and $z_{u,i} = [\mathbf{p''_u}, \mathbf{q''_i}]$. Ideally, a desirable disentangling would make $x_{u,i}$ and $z_{u,i}$ independent. Meanwhile, $x_{u,i}$ should have the ability to predict $r_{u,i}$, whereas the concatenation of $x_{u,i}$ and $z_{u,i}$ should have the ability to predict $o_{u,i}$. Motivated by this, the proposed method adopts such disentanglement (DT for short) to the previous vanilla IPS and vanilla DR estimators, named DT-IPS and DT-DR, respectively.

For DT-IPS and DT-DR, we propose to use multi-task learning to ensure the performance of various tasks. The prediction model of DT-IPS is trained by minimizing the following loss function

$$\begin{aligned}\mathcal{L}_{DT-IPS}(\mathbf{P}, \mathbf{Q}; \theta_r, \theta_o) &= \underbrace{\mathcal{L}_{IPS}(\mathbf{P}', \mathbf{Q}'; \theta_r)}_{\text{Vanilla IPS loss}} \\ &+ \alpha \underbrace{\mathcal{L}_O(\mathbf{P}, \mathbf{Q}; \theta_o)}_{\text{Propensity loss}} + \beta \underbrace{\left(\left\|\mathbf{P}'^{\top}\mathbf{P}''\right\|_{\mathrm{F}}^2 + \left\|\mathbf{Q}'^{\top}\mathbf{Q}''\right\|_{\mathrm{F}}^2\right)}_{\text{Disentangling loss}} \\ &+ \gamma \underbrace{\left(\left\|\mathbf{P}'\mathbf{Q}'^{\top}\right\|_{\mathrm{F}}^2 + \left\|\mathbf{P}''\mathbf{Q}''^{\top}\right\|_{\mathrm{F}}^2\right)}_{\text{Regularization loss}},\end{aligned}$$

where $\mathbf{P} = [\mathbf{P}', \mathbf{P}''] \in \mathbb{R}^{|U| \times K}$, $\mathbf{Q} = [\mathbf{Q}', \mathbf{Q}''] \in \mathbb{R}^{|I| \times K}$, $\mathbf{P}' \in \mathbb{R}^{|U| \times A}$, $\mathbf{P}'' \in \mathbb{R}^{|U| \times (K-A)}$, $\mathbf{Q}' \in \mathbb{R}^{|I| \times A}$, $\mathbf{Q}'' \in \mathbb{R}^{|I| \times (K-A)}$, and A is the hyper-parameter to control the disentangling dimensions of $x_{u,i}$ and $z_{u,i}$.

As shown in Figures 1(d) and 2, we use the concatenation of $x_{u,i}$ and $z_{u,i}$, i.e., $[\mathbf{p_u}, \mathbf{q_i}]$, to predict $o_{u,i}$. Therefore, the propensity model is trained by minimizing the following cross-entropy loss function

$$\begin{aligned}&\mathcal{L}_O(\mathbf{P}, \mathbf{Q}; \theta_o) \\ &= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \Big[- o_{u,i} \cdot \log \hat{p}_{u,i} - (1 - o_{u,i}) \cdot \log(1 - \hat{p}_{u,i}) \Big],\end{aligned}$$

where $\hat{p}_{u,i} = m(\mathbf{p_u}, \mathbf{q_i}; \theta_o)$ is an estimate of $\mathbb{P}(o_{u,i} = 1 \mid x_{u,i}, r_{u,i})$. Meanwhile, we use $x_{u,i} = [\mathbf{p'_u}, \mathbf{q'_i}]$ to predict the rating $r_{u,i}$. Thus, the $\mathcal{L}_{IPS}(\mathbf{P}', \mathbf{Q}'; \theta_r)$ is defined as

$$\mathcal{L}_{IPS}(\mathbf{P}', \mathbf{Q}'; \theta_r) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}},$$

where $e_{u,i} = (r_{u,i} - f(x'_{u,i}; \theta_r))^2 = (r_{u,i} - f(\mathbf{p'_u}, \mathbf{q'_i}; \theta_r))^2$.

The four Frobenius norms (F norms), i.e., $\left\|\mathbf{P}'^{\top}\mathbf{P}''\right\|_{\mathrm{F}}^2$, $\left\|\mathbf{Q}'^{\top}\mathbf{Q}''\right\|_{\mathrm{F}}^2$, $\left\|\mathbf{P}'\mathbf{Q}'^{\top}\right\|_{\mathrm{F}}^2$, and $\left\|\mathbf{P}''\mathbf{Q}''^{\top}\right\|_{\mathrm{F}}^2$, play an important role in the proposed DT-IPS loss. The core part of the disentangling is to ensure the $x_{u,i}$ and $z_{u,i}$ to be independent. Specifically, we need both $\mathbf{P}'$ and $\mathbf{P}''$, $\mathbf{Q}'$ and $\mathbf{Q}''$ independent. However, we cannot use the inner product or cosine similarity as the constraint because the two vectors have different lengths when $A \neq \frac{K}{2}$. Therefore, we follow [57] to use the outer product instead. When the F norm of the outer product is zero, every element-wise product is zero, which ensures the independence of the two vectors. In addition, the last two terms can prevent overfitting and it will make the contribution of all features more dispersed, instead of leaving some features completely dominant.

For the DT-DR method, we need the error imputation model $\hat{e}_{u,i} = g(\mathbf{U}, \mathbf{V}; \theta_e)$ to estimate the prediction loss $e_{u,i}$, where U is the user feature embedding and V is the item feature

embedding. We substitute $\mathcal{L}_{IPS}$ in the DT-IPS loss by the following DR loss to obtain the DT-DR loss

$$\mathcal{L}_{DR}(\mathbf{P}', \mathbf{Q}', \mathbf{U}, \mathbf{V}) = \mathcal{L}_{DR}^{err}(\mathbf{P}', \mathbf{Q}'; \theta_r) + \mathcal{L}_{DR}^{imp}(\mathbf{U}, \mathbf{V}; \theta_e),$$

where the DR loss for debiased learning is on $\mathbf{P}'$ and $\mathbf{Q}'$

$$\mathcal{L}_{DR}^{err}(\mathbf{P}', \mathbf{Q}'; \theta_r) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \left(\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}} \right),$$

and the error imputation model loss is on $\mathbf{U}$ and $\mathbf{V}$

$$\mathcal{L}_{DR}^{imp}(\mathbf{U}, \mathbf{V}; \theta_e) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})^2}{\hat{p}_{u,i}}.$$

We summarize the propensity and prediction collaborative filtering with the orthogonal constraints in the proposed disentanglement approach in Figure 2.

C. Training Efficiency

Table II compares the training efficiency of various debiasing methods in terms of the parameter size and training loss, as the main factors affecting the training efficiency. For the parameter size, we compared the embedding size and hidden layer parameter size, respectively, by taking the ESMM [19], which is widely used in industrial applications, as a benchmark. Compared to vanilla IPS and DR methods with 2x and 3x embedding sizes, respectively, Multi-IPS/DR [16] and ESCM^2-IPS/DR [18] have the same embedding size as the ESMM due to the inclusion of an embedding lookup table to share embeddings among the propensity, computation, and prediction models. For our proposed DT-IPS and DT-DR, as shown in Figure 2, the embedding of the prediction model is contained in that of the propensity model, whereas the embedding of the imputation model of DT-DR incurs additional overhead. As a result, DT-IPS and DT-DR have the same and 2x the embedding size compared to ESMM, respectively. In addition, the IPS and DR methods have the same and 2x the hidden layer parameter size compared to ESMM, respectively, due to the DR methods requiring additional parameters from the imputation model, leading to smaller bias in practice.

Moreover, the number and type of training losses also significantly affect the number and time of backpropagation. Compared to Multi-IPS and Multi-DR, which use separate IPS and DR losses to train all models, ESCM^2-IPS and ESCM^2-DR further use propensity and post-view clickthrough&conversion rate (CTCVR) losses to enhance the debiasing performance. Instead, our DT-IPS and DT-DR use propensity and disentangling losses, where the latter requires the computation of the F norm for a large user-item matrix, which may result in a slight additional time overhead.

TABLE II: Comparison of parameter numbers in terms of embedding size, hidden layer size, and training loss.

	Parameter size		Training loss		
Method	Embedding	Hidden layer	Propensity	CTCVR	Disentangle
ESMM [19]	1×	1×	✓	✓	×
IPS [6]	2×	1×	✓	×	×
Multi-IPS [16]	1×	1×	×	×	×
ESCM^2-IPS [18]	1×	1×	✓	✓	×
DT-IPS (ours)	1×	1×	✓	×	✓
DR [15], [17], [58]	3×	1.5×	×	×	×
Multi-DR [16]	1×	1.5×	×	×	×
ESCM^2-DR [18]	1×	1.5×	✓	✓	×
DT-DR (ours)	2×	1.5×	✓	×	✓

V. Semi-synthetic Experiments

Experiment Setup. Following the previous studies [6], [15], [17], [58], we conduct semi-synthetic experiments on the MovieLens 100K (ML-100K)[1] dataset, which contains 943 users and 1,682 items with 100,000 observed MNAR five-scale ratings. The main purpose is to verify the effectiveness of the proposed method when the rating $r_{u,i}$ affects the observation $o_{u,i}$. The dataset pre-processing procedure and the experiment details are shown below:

Step 1. Following the previous studies [6], [15], [17], we generate the conversion probabilities for all user-item pairs using the matrix factorization (MF) [59]. Specifically, we first minimize the mean squared loss on the observed ratings to train the MF model and generate a rating for each user-item pair. Then we clip the generated rating to $[0, 5]$, denoted as $\gamma_{u,i}$, and use the following formula for standardization to ensure the generated conversion probability $\eta_{u,i} \in [0, 1]$ for all user-item pairs:

$$\eta_{u,i} = \epsilon + (1 - \epsilon)\frac{\gamma_{u,i} - \gamma_{\min}}{\gamma_{\max} - \gamma_{\min}}, \quad \forall (u, i) \in \mathcal{D}, \tag{11}$$

where $\gamma_{\min}$ and $\gamma_{\max}$ are the minimum and maximum value of the generated ratings after clipping, respectively, and $\epsilon \in [0, 1]$ is a pre-specified hyper-parameter to control the noise.

Step 2. To make the conversion probability affect the observed probability, we set the observed probability as $p_{u,i} = (2^{\eta_{u,i}} - 1)^{\rho}$ for all user-item pairs, where ρ is the hyper-parameter to control the observed data sparsity and the degree of correlation between $\eta_{u,i}$ and $p_{u,i}$.

Step 3. We sample the binarized conversion result $r_{u,i}$ and the observed indicator $o_{u,i}$ by the following Bernoulli distribution:

$$r_{u,i} \sim \text{Bern}(\eta_{u,i}), \quad o_{u,i} \sim \text{Bern}(p_{u,i}), \quad \forall (u, i) \in \mathcal{D}.$$

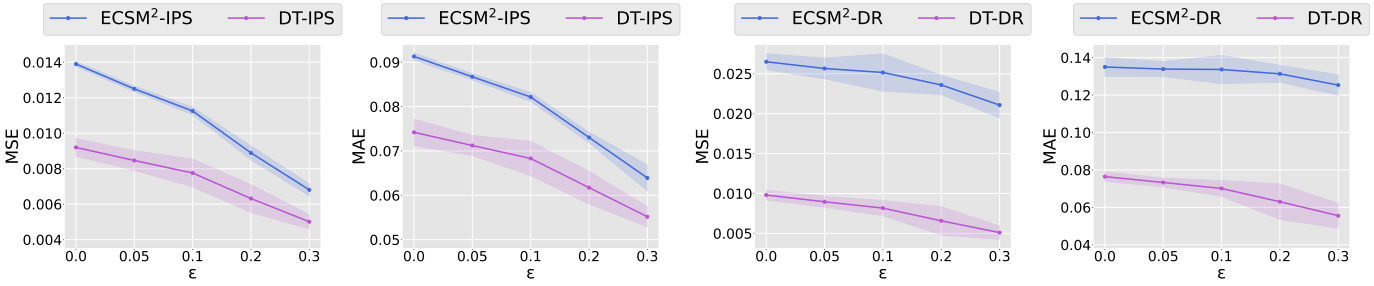
Then we use the generated $r_{u,i}$ and $o_{u,i}$ to train the prediction model. Since we directly use $\eta_{u,i}$ to generate $p_{u,i}$, so there is a strong correlation between the observed indicators $o_{u,i}$ and the conversion results $r_{u,i}$.

Performance Analysis. To verify the effectiveness of the proposed method, Table III shows the performance of the proposed method and the baseline methods under varying ρ. Three metrics are used to evaluate the performance: mean square error (MSE), mean absolute error (MAE), and NDCG@50 (N@50). First, the proposed method achieves overall better performance when ρ is larger. This is because the $r_{u,i}$ has a larger influence on $o_{u,i}$ with a larger ρ, which strengthens the necessity of disentangling. Meanwhile, the proposed

[1]https://grouplens.org/datasets/movielens/100k/

TABLE III: Performance comparison on the ML-100K with varying ρ. The best result is bolded and the second is underlined.

Metrics	MSE					MAE					N@50				
ρ	0.5	0.75	1	1.25	1.5	0.5	0.75	1	1.25	1.5	0.5	0.75	1	1.25	1.5
MF	0.0154	0.0148	0.0142	0.0138	0.0134	0.1033	0.1005	0.0981	0.0960	0.0942	0.8015	0.8015	0.8016	0.8016	0.8015
IPS	0.0124	0.0126	0.0128	0.0129	0.0130	0.0863	0.0870	0.0876	0.0881	0.0885	0.8016	0.8017	0.8017	0.8017	0.8017
DR	0.0123	0.0124	0.0126	0.0128	0.0130	0.0859	0.0865	0.0871	0.0877	0.0886	0.8017	0.8016	0.8016	0.8016	0.8016
Multi-IPS	0.0139	0.0139	0.0139	0.0139	0.0138	0.0916	0.0917	0.0917	0.0916	0.0917	0.8017	0.8016	0.8016	0.8016	0.8016
Multi-DR	0.0243	0.0308	0.0361	0.0402	0.0435	0.1231	0.1383	0.1493	0.1577	0.1627	0.8014	0.8009	0.8009	0.8014	0.8015
ESCM^2-IPS	0.0139	0.0139	0.0139	0.0139	0.0139	0.0913	0.0913	0.0913	0.0913	0.0913	0.8017	0.8016	0.8017	0.8017	0.8017
ESCM^2-DR	0.0173	0.0214	0.0265	0.0325	0.0394	0.1044	0.1187	0.1350	0.1531	0.1719	0.8016	0.8015	0.8014	0.8015	0.8014
DT-IPS	0.0193	0.0109	**0.0092**	0.0103	0.0121	0.1155	0.0815	**0.0742**	0.0795	0.0874	**0.8022**	**0.8021**	**0.8020**	**0.8019**	**0.8019**
DT-DR	**0.0118**	**0.0103**	0.0098	**0.0097**	**0.0097**	**0.0845**	**0.0784**	0.0764	**0.0762**	**0.0766**	0.8019	0.8020	0.8018	0.8017	0.8017



(a) MSE of IPS (b) MAE of IPS (c) MSE of DR (d) MAE of DR

Fig. 3: MSE and MAE of IPS and DR estimators with varying noise hyper-parameter ϵ.

method outperforms the baseline methods with varying hyper-parameter ρ, which further verifies the effectiveness of the proposed method.

Moreover, we vary hyper-parameter ϵ in equation (11) to further validate the effectiveness of the proposed method under different data generation mechanisms. The experimental results are shown in Figure 3. First, as ϵ increases, the MSE and MAE of all methods become lower, which is because the heterogeneity between the user-item pairs is reduced. Second, the proposed method stably and statistically significantly outperforms the baselines under varying ϵ, demonstrating the effectiveness of our method under different settings.

VI. Real-World Experiments

Datasets. We conduct the real-world experiments on three widely used real-world datasets including Coat[2], Yahoo[3] and a large-scale industrial dataset KuaiRec[4] [62]. Coat contains ratings from 290 users to 300 items. Each user evaluates 24 items, containing 6,960 MNAR ratings in total. Meanwhile, each user evaluates 16 randomly selected items, which results in 4,640 MAR ratings. Yahoo contains a total of 311,704 MNAR and 54,000 MAR ratings from 15,400 users to 1,000 items. KuaiRec contains 12,530,806 MNAR watching ratio from 7,176 users to 10,728 videos, as well as a separate MAR test set where the users are asked to rate all test items. For the first two five-scale datasets, the ratings less than three are clipped to 0, otherwise to 1. For KuaiRec, the video records less than one are clipped to 0, otherwise to 1.

Baselines. In pursuit of a comprehensive comparison, we use matrix factorization (MF) [59] as our base model, and we have compared the proposed method with the following debiased baseline methods, including IPS [6], DR [58], TDR [50], and Stable-DR [49]; joint learning methods such as DR-JL [15], MRDR-JL [17], DR-BIAS [53], DR-MSE [53], MR [52], and TDR-JL [50]; multi-task learning approaches including Multi-IPS [16], Multi-DR [16], ESMM [19], ESCM^2-IPS [18], ESCM^2-DR [18], IPS-V2 [61], and DR-V2 [61]; the information bottleneck based methods: CVIB [60] and DIB [21]. The details of each method are shown below:

- **MF** [59]: MF is the traditional method that factorizes the user-item matrix into user and item latent vectors and uses the inner product between the latent vector to impute missing values and make predictions.
- **IPS** [6]: IPS utilizes the propensity to reweight the observed prediction error to obtain unbiased estimation.
- **DR** [58]: DR combines the merits of inverse propensity weighting and error imputation to achieve the double robustness property to reduce estimation error.
- **TDR** [50]: TDR is one of the state-of-the-art DR method that learns imputation with a parameterized imputation model and a nonparametric boosting strategy.
- **Stable-DR** [49]: Stable-DR has a weaker reliance on extrapolation and has bounded bias, variance, and gen-

[2]https://www.cs.cornell.edu/~schnabts/mnar/
[3]http://webscope.sandbox.yahoo.com/
[4]https://github.com/chongminggao/KuaiRec

TABLE IV: Performance comparison on COAT, YAHOO, and KUAIREC. The two best results are bolded, and the best baseline results are underlined. $*$ means statistically significant (p-value ≤ 0.05) using the paired t-test compared with the best baseline.

	COAT			YAHOO			KUAIREC		
Method	AUC	N@5	R@5	AUC	N@5	R@5	AUC	N@50	R@50
MF [59]	$0.696_{\pm 0.007}$	$0.641_{\pm 0.007}$	$0.453_{\pm 0.010}$	$0.680_{\pm 0.002}$	$0.648_{\pm 0.002}$	$0.415_{\pm 0.002}$	$0.741_{\pm 0.002}$	$0.725_{\pm 0.001}$	$0.800_{\pm 0.002}$
CVIB [60]	$0.700_{\pm 0.006}$	$0.643_{\pm 0.011}$	$0.456_{\pm 0.011}$	$0.678_{\pm 0.001}$	$0.634_{\pm 0.001}$	$0.393_{\pm 0.002}$	$0.744_{\pm 0.014}$	$0.737_{\pm 0.015}$	$0.817_{\pm 0.016}$
DIB [21]	$0.721_{\pm 0.004}$	$0.645_{\pm 0.010}$	$0.455_{\pm 0.010}$	$0.689_{\pm 0.001}$	$0.639_{\pm 0.001}$	$0.400_{\pm 0.001}$	$0.781_{\pm 0.008}$	$0.792_{\pm 0.007}$	$0.845_{\pm 0.009}$
IPS [6]	$0.698_{\pm 0.006}$	$0.640_{\pm 0.011}$	$0.455_{\pm 0.011}$	$0.680_{\pm 0.002}$	$0.659_{\pm 0.002}$	$0.427_{\pm 0.002}$	$0.730_{\pm 0.001}$	$0.714_{\pm 0.002}$	$0.797_{\pm 0.002}$
DR [58]	$0.700_{\pm 0.006}$	$0.658_{\pm 0.009}$	$0.462_{\pm 0.010}$	$0.646_{\pm 0.003}$	$0.676_{\pm 0.002}$	$0.435_{\pm 0.001}$	$0.709_{\pm 0.002}$	$0.680_{\pm 0.002}$	$0.782_{\pm 0.002}$
DR-JL [15]	$0.717_{\pm 0.006}$	$0.650_{\pm 0.005}$	$0.451_{\pm 0.006}$	$0.688_{\pm 0.001}$	$0.648_{\pm 0.001}$	$0.409_{\pm 0.001}$	$0.776_{\pm 0.006}$	$0.788_{\pm 0.006}$	$0.841_{\pm 0.009}$
MRDR-JL [17]	$0.718_{\pm 0.004}$	$0.666_{\pm 0.007}$	$0.451_{\pm 0.006}$	$0.688_{\pm 0.001}$	$0.647_{\pm 0.002}$	$0.408_{\pm 0.002}$	$0.780_{\pm 0.005}$	$0.792_{\pm 0.005}$	$0.840_{\pm 0.009}$
DR-BIAS [53]	$0.714_{\pm 0.008}$	$0.641_{\pm 0.010}$	$0.450_{\pm 0.011}$	$0.686_{\pm 0.001}$	$0.653_{\pm 0.001}$	$0.430_{\pm 0.002}$	$0.754_{\pm 0.002}$	$0.740_{\pm 0.002}$	$0.816_{\pm 0.003}$
DR-MSE [53]	$0.720_{\pm 0.004}$	$0.641_{\pm 0.009}$	$0.449_{\pm 0.009}$	$0.687_{\pm 0.002}$	$0.659_{\pm 0.002}$	$0.432_{\pm 0.002}$	$0.754_{\pm 0.002}$	$0.742_{\pm 0.003}$	$0.817_{\pm 0.002}$
MR [52]	$0.727_{\pm 0.004}$	$0.654_{\pm 0.009}$	$0.469_{\pm 0.009}$	$\underline{0.700_{\pm 0.002}}$	$0.682_{\pm 0.002}$	$0.449_{\pm 0.002}$	$\underline{0.792_{\pm 0.002}}$	$0.799_{\pm 0.003}$	$0.843_{\pm 0.002}$
TDR [50]	$0.712_{\pm 0.008}$	$0.644_{\pm 0.010}$	$0.451_{\pm 0.011}$	$0.690_{\pm 0.002}$	$0.691_{\pm 0.003}$	$0.457_{\pm 0.003}$	$0.753_{\pm 0.002}$	$0.722_{\pm 0.002}$	$0.820_{\pm 0.001}$
TDR-JL [50]	$0.719_{\pm 0.008}$	$0.623_{\pm 0.011}$	$0.448_{\pm 0.011}$	$0.688_{\pm 0.001}$	$0.649_{\pm 0.001}$	$0.413_{\pm 0.002}$	$0.778_{\pm 0.005}$	$0.790_{\pm 0.007}$	$0.837_{\pm 0.011}$
Stable-DR [49]	$0.717_{\pm 0.009}$	$0.613_{\pm 0.015}$	$0.426_{\pm 0.011}$	$0.689_{\pm 0.001}$	$0.651_{\pm 0.003}$	$0.414_{\pm 0.003}$	$0.761_{\pm 0.004}$	$0.786_{\pm 0.003}$	$0.834_{\pm 0.006}$
Multi-IPS [16]	$0.720_{\pm 0.006}$	$0.600_{\pm 0.012}$	$0.437_{\pm 0.008}$	$0.683_{\pm 0.002}$	$0.656_{\pm 0.002}$	$0.435_{\pm 0.003}$	$0.764_{\pm 0.019}$	$0.758_{\pm 0.020}$	$0.837_{\pm 0.020}$
Multi-DR [16]	$0.721_{\pm 0.005}$	$0.650_{\pm 0.005}$	$0.459_{\pm 0.008}$	$0.686_{\pm 0.003}$	$0.660_{\pm 0.002}$	$0.420_{\pm 0.002}$	$0.760_{\pm 0.014}$	$0.767_{\pm 0.012}$	$0.826_{\pm 0.012}$
ESMM [19]	$0.708_{\pm 0.008}$	$\underline{0.681_{\pm 0.009}}$	$\underline{0.479_{\pm 0.010}}$	$0.668_{\pm 0.002}$	$\underline{0.730_{\pm 0.001}}$	$\underline{0.495_{\pm 0.002}}$	$0.743_{\pm 0.004}$	$0.788_{\pm 0.004}$	$0.836_{\pm 0.003}$
ESCM^2-IPS [18]	$0.714_{\pm 0.011}$	$0.654_{\pm 0.025}$	$0.469_{\pm 0.018}$	$0.688_{\pm 0.001}$	$0.673_{\pm 0.003}$	$0.445_{\pm 0.004}$	$0.781_{\pm 0.001}$	$0.770_{\pm 0.002}$	$0.836_{\pm 0.002}$
ESCM^2-DR [18]	$0.725_{\pm 0.006}$	$0.645_{\pm 0.012}$	$0.450_{\pm 0.007}$	$0.693_{\pm 0.002}$	$0.688_{\pm 0.003}$	$0.449_{\pm 0.003}$	$0.791_{\pm 0.001}$	$\mathbf{0.799}_{\pm 0.001}$	$\underline{0.854_{\pm 0.001}}$
IPS-V2 [61]	$0.724_{\pm 0.004}$	$0.629_{\pm 0.009}$	$0.455_{\pm 0.009}$	$0.685_{\pm 0.002}$	$0.646_{\pm 0.002}$	$0.457_{\pm 0.002}$	$0.774_{\pm 0.002}$	$\underline{0.768_{\pm 0.003}}$	$\underline{0.838_{\pm 0.002}}$
DR-V2 [61]	$\mathbf{0.734}_{\pm 0.004}$	$0.646_{\pm 0.009}$	$0.452_{\pm 0.009}$	$0.690_{\pm 0.002}$	$0.660_{\pm 0.002}$	$0.475_{\pm 0.002}$	$0.783_{\pm 0.002}$	$0.788_{\pm 0.003}$	$0.842_{\pm 0.002}$
DT-IPS (ours)	$0.733_{\pm 0.005}$	$\mathbf{0.715}^{*}_{\pm 0.014}$	$\mathbf{0.503}^{*}_{\pm 0.011}$	$\mathbf{0.715}^{*}_{\pm 0.003}$	$\mathbf{0.740}^{*}_{\pm 0.002}$	$\mathbf{0.503}^{*}_{\pm 0.002}$	$\mathbf{0.806}^{*}_{\pm 0.001}$	$\mathbf{0.813}^{*}_{\pm 0.001}$	$\mathbf{0.869}^{*}_{\pm 0.001}$
DT-DR (ours)	$\mathbf{0.738}^{*}_{\pm 0.008}$	$\mathbf{0.705}^{*}_{\pm 0.011}$	$\mathbf{0.509}^{*}_{\pm 0.014}$	$\mathbf{0.707}^{*}_{\pm 0.002}$	$\mathbf{0.765}^{*}_{\pm 0.001}$	$\mathbf{0.526}^{*}_{\pm 0.001}$	$\mathbf{0.803}^{*}_{\pm 0.001}$	$0.797_{\pm 0.001}$	$\mathbf{0.859}^{*}_{\pm 0.001}$

eralization error bound simultaneously.

- **DR-JL** [15]: DR-JL joint learns the rating prediction model and error imputation model to achieve promising prediction performance guarantees based on vanilla DR.
- **MRDR-JL** [17]: MRDR-JL further reduces the DR estimator's variance while retaining its double robustness.
- **DR-BIAS** [53]: DR-BIAS is a DR-based method that further controls the bias of the DR loss.
- **DR-MSE** [53]: DR-MSE is an extension of the vanilla DR that balances the bias and variance flexibly.
- **MR** [52]: MR is an enhanced version of DR, which combines multiple propensity models and imputation models, and greatly relaxes the unbiased condition.
- **TDR-JL** [50]: TDR-JL adopts joint learning on TDR.
- **Multi-IPS** [16]: Multi-IPS utilizes the multi-task learning on vanilla IPS estimator to reduce the bias.
- **Multi-DR** [16]: Similar to Multi-IPS, Multi-DR uses the multi-task learning on vanilla DR estimator.
- **ESMM** [19]: ESMM trains the prediction model on the entire space to align the observed and target population.
- **ESCM^2-IPS** [18]: ESCM^2-IPS employs a counterfactual risk minimizer as a regularizer in ESMM to address biases on the prediction inaccuracy estimation.
- **ESCM^2-DR** [18]: ESCM^2-DR adds an imputation model based on ESCM^2-IPS to guarantee the double robustness.
- **IPS-V2** [61]: IPS-V2 utilizes the balancing propensities to achieve smaller variance compared to the IPS.
- **DR-V2** [61]: DR-V2 adds an error imputation model based on IPS-V2 to guarantee the double robustness.
- **CVIB** [60]: CVIB derives a contrastive information loss and an output confidence penalty to facilitate balanced learning between the factual and counterfactual domains.
- **DIB** [21]: DIB constrains the model to learn a biased embedding vector with independent biased and unbiased components in the training phase and uses only the unbiased component in the test phase.

Experimental Protocols. Following previous studies [6], [15], [53], we adopt three widely used evaluation metrics, AUC, NDCG@K (N@K) and Recall@K (R@K), where $K = 5$ for **Coat** and **Music**, while $K = 50$ for **KuaiRec**. NDCG@K evaluates the quality of recommendations by considering the importance of each item's position based on discounted gains:

$$DCG_u@K = \sum_{i \in D^u_{\text{test}}} \frac{I(\hat{z}_{u,i} \leq K)}{\log(\hat{z}_{u,i} + 1)},$$

$$NDCG@K = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{DCG_u@K}{IDCG_u@K},$$

where D^u_{test} denotes all the items rated by the user u in test data, $\hat{z}_{u,i}$ represents the ranking of item i in the recommended list for user u, and IDCG represents the best possible DCG. In addition, the formula of Recall@K is as follows:

$$Recall_u@K = \frac{\sum_{i \in D^u_{\text{test}}} I(\hat{z}_{u,i} \leq K)}{\min(K, |D^u_{\text{test}}|)},$$

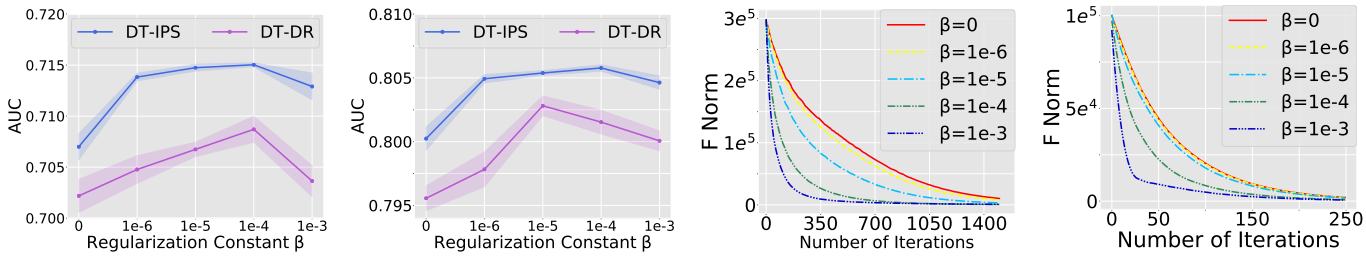
$$Recall@K = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} Recall_u@K.$$

Experimental Details. Throughout the parameter-tuning process, all methods are implemented on PyTorch with Adam as the optimizer[5]. We tune the learning rate

[5]In all experiments, we use Tesla T4 GPU as the computational resource.

TABLE V: Ablation studies in terms of the training losses. The best results are bolded, and the second best result is underlined.

	Training loss		COAT			YAHOO			KUAIREC		
Method	β	γ	AUC	N@5	R@5	AUC	N@5	R@5	AUC	N@50	R@50
DT-IPS	×	×	0.717	0.678	0.483	0.700	0.726	0.493	0.768	0.749	0.831
	×	✓	0.722	0.685	0.492	0.707	0.732	0.499	0.800	0.809	0.864
	✓	×	0.726	0.689	0.488	0.711	0.738	0.501	0.802	0.810	0.863
	✓	✓	**0.733**	**0.715**	**0.503**	**0.715**	**0.740**	**0.503**	**0.806**	**0.813**	**0.869**
DT-DR	×	×	0.703	0.632	0.450	0.675	0.742	0.503	0.783	0.774	0.846
	×	✓	0.723	0.671	0.483	0.702	0.763	0.523	0.796	0.790	0.851
	✓	×	0.727	0.683	0.506	0.705	0.762	0.524	0.799	0.792	0.857
	✓	✓	**0.738**	**0.705**	**0.509**	**0.707**	**0.765**	**0.526**	**0.803**	**0.797**	**0.859**



(a) AUC on YAHOO (b) AUC on KUAIREC (c) F norm on Yahoo (d) F norm on KUAIREC

Fig. 4: Sensitivity analysis of hyper-parameter β on the YAHOO and KUAIREC datasets.

in $\{0.005, 0.01, 0.05, 0.1\}$, batch size in $\{32, 64, 128, 256\}$ for COAT and $\{1024, 2048, 4096, 8192\}$ for YAHOO and KUAIREC, and embedding dimension in $\{2, 4, 8, 16, 32, 64\}$ for COAT and $\{8, 16, 32, 64, 128, 256\}$ for YAHOO and KUAIREC. For our method, we tune the weights β and γ in $\{1e-6, 5e-6, 1e-5, ..., 5e-2, 1e-1\}$.

A. Performance Comparison

We train the prediction models with biased ratings and evaluate them with unbiased ratings on three widely used real-world datasets, COAT, YAHOO, and a large-scale industrial dataset KUAIREC. The results are shown in Table IV, and we have the following findings. For the COAT and YAHOO datasets, almost all debiasing methods exhibit better performance than the base model, which indicates the necessity of modeling data-missing mechanisms in MNAR settings. Next, DR-based methods outperform IPS-based methods overall, and the multi-task learning method ESMM achieves the most competitive performance among all baselines, which is attributed to the fact that ESMM models rating prediction task over the entire space to eliminate the gap between observed and target population. Moreover, the proposed method stably outperforms baseline methods on all datasets, due to the effectiveness of the disentangling and multi-task learning to accurately estimate the MNAR propensities.

For KUAIREC, we find that two DR-enhanced methods, MR and ESCM^2-DR, achieve the most competitive performance among the baselines. This is attributed to the relaxation of the unbiasedness condition by MR using multiple candidate propensity and imputation models, as well as the counterfactual risk minimizer on the entire space in ESCM^2-DR. Besides, despite the proposed DT-DR being slightly underperformed by ESCM^2-DR in terms of NDCG@50, DT-DR significantly outperforms ESCM^2-DR in terms of both AUC and Recall@50, leading to the optimal overall results. This further provides empirical evidence that the real-world data are MNAR and verifies the effectiveness of the proposed method.

B. Ablation Study

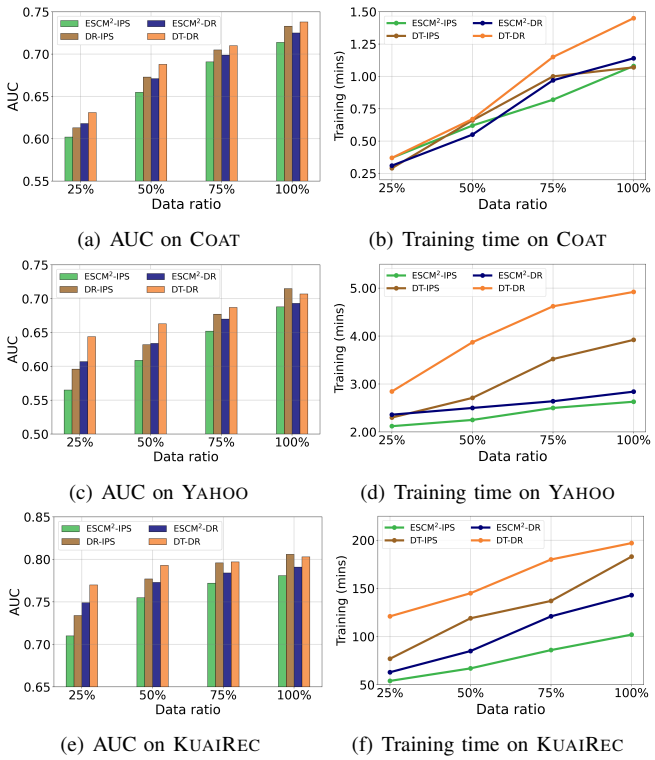
The disentangling loss and regularization loss play an important role in our method. To investigate the reason for the performance improvement of our method, Table V shows the ablation experiment results on all three datasets. First, the proposed method with both losses always achieves the best performance and the method with only disentangling loss achieves secondary performance. It is because the missing regularization term simply controls the F norm to prevent overfitting, which is not the main focus in data MNAR. Meanwhile, the method with only the regularization term performs worse than the former two because there is no guarantee of the disentangling quality. Finally, without both the regularization term and disentangling term will lead to the worst performance, which further verifies the effectiveness and rationality of our method.

C. Sensitivity Analysis

The hyper-parameter β controls the degree of independence of the disentangled two parts, that is, the user embedding $\mathbf{p}'_\mathbf{u}$ and $\mathbf{p}''_\mathbf{u}$ and the item embedding $\mathbf{q}'_\mathbf{i}$ and $\mathbf{q}''_\mathbf{i}$. Figures 4(a) and 4(b) show the sensitivity analysis results of hyper-parameter β on the prediction performance using the YAHOO

TABLE VI: Comparison of parameter numbers, training time (minutes), and inference time (milliseconds per sample).

	COAT			YAHOO			KUAIREC		
Method	Parameters	Training	Inference	Parameters	Training	Inference	Parameters	Training	Inference
ESMM [19]	2.30×10^4	0.40	1.06	2.12×10^6	2.51	1.61	1.02×10^7	66.53	0.66
IPS [6]	3.77×10^4	0.53	0.92	4.19×10^6	1.92	1.07	1.83×10^7	50.20	0.78
Multi-IPS [16]	2.30×10^4	1.13	0.94	2.12×10^6	1.30	1.70	1.02×10^7	82.08	2.67
ESCM^2-IPS [18]	2.30×10^4	1.08	1.34	2.12×10^6	2.63	1.86	1.02×10^7	102.22	1.39
DT-IPS (ours)	1.88×10^4	1.07	1.18	2.09×10^6	3.92	1.01	0.92×10^7	183.76	0.92
DR-JL [15]	5.66×10^4	1.21	1.40	6.29×10^6	3.10	1.91	2.75×10^7	139.37	0.74
Multi-DR [16]	2.51×10^4	1.20	1.47	2.14×10^6	3.19	2.15	1.07×10^7	117.13	1.30
ESCM^2-DR [18]	2.51×10^4	1.14	1.01	2.14×10^6	2.84	2.12	1.07×10^7	143.34	1.71
DT-DR (ours)	3.77×10^4	1.45	1.45	4.19×10^6	4.92	0.73	1.83×10^7	197.48	0.94



(a) AUC on COAT (b) Training time on COAT

(c) AUC on YAHOO (d) Training time on YAHOO

(e) AUC on KUAIREC (f) Training time on KUAIREC

Fig. 5: Effect of data sparsity on AUC and training time.

and KUAIREC datasets. We find that the moderate choices of β such as $1e-4$ or $1e-5$ lead to optimal performance. On one hand, there will be no guarantee of the disentangling quality when β goes small or reaches at 0, which harms the prediction performance. On the other hand, when β goes large, the performance also gets worse because putting too much attention on the disentangling will make the optimization for other tasks such as rating prediction or propensity learning insufficient, resulting in larger value of other losses, which also harms the performance. In addition, Figures 4(c) and 4(d) show the effect of the hyper-parameter β on the disentangling loss scale, which is the summation of two F norms in the disentangling loss on the YAHOO and KUAIREC datasets. As the number of iterations increases, the value of the disentangling loss decreases. Meanwhile, as the increasing of β, the value of the disentangling loss converges at a faster speed.

D. Model Training and Inference Time

Algorithm complexity and algorithm runtime are very important in the database field. We report the results on the number of parameters, training time, and inference time in Table VI. First, note that many previous methods based on the embedding parameter sharing mechanism cannot be used directly with MF as the base model (which would result in an identical prediction model, propensity model, and imputation model). Therefore, we use a shallow MLP to implement these methods after the embedding layer. The DT-IPS method has the fewest parameters among the IPS-based methods and DT-DR has much fewer parameters compared with the DR-JL method, which demonstrates the parameter efficiency of our methods so that DT-IPS and DT-DR consume less storage space. In addition, the running time is not much different from the baseline methods on smaller datasets (COAT and YAHOO), but takes 1.5-2 times longer than the previous multi-task learning methods on a larger industrial dataset (KUAIREC). The reason is that for larger datasets, the multiplication of matrices consumes more time when computing the norms. Notably, the debiasing prediction performance of the proposed methods significantly outperform the baseline methods in almost all the cases as shown in Table IV.

E. Scalability

Data sparsity is also a widespread concern in the database field, which motivates us to analyze the recommendation performance and runtime of our method and baselines under different data sparsity. The results are shown in Figure 5. The runtimes of our method under different data sparsity are all within 2 times compared to the baseline methods, and the runtimes of our method are almost the same as the baseline methods on the COAT dataset. Meanwhile, our method significantly outperforms the baseline methods, which further proves the effectiveness of our method.

VII. Related Work

Missing Not At Random (MNAR). MNAR problem is ubiquitous in many fields such as information retrieval [63], biology [28], and social science [64]. Identifying the model parameters with data MNAR is challenging because the full data is only partially observed. Several previous studies considered the identifiability problem for some specific parametric and semi-parametric models, e.g., probit regression, normal mixture, and t mixture model with non-monotone missing mechanisms [55], [65]–[68]. Nevertheless, many parametric models of interest, e.g., logistic regression, are not identifiable under MNAR without imposing additional assumptions [67].

Several studies explored the identifiability results using additional informative variables. For example, [69]–[71] discussed the conditions for identifiability by leveraging an instrumental variable that only affects the missing mechanism but not the outcome. In addition, [67] and [72] established the identifiability results by a shadow variable. Such a variable is associated with the potentially unobserved variable conditional on the observed data, but independent of the missing mechanism conditional on both the observed and missing values.

Even after the establishment of the identifiability results, another challenge arises with data MNAR in the parameter estimation process, which is the model misspecification problem. To address such issues, likelihood-based inference [68], [73], imputation-based methods [74], inverse probability weighting [55], and doubly robust estimation [71] have been developed. These existing estimation methods necessitate the correct model specification of either the propensity score or the imputation model. However, bias may arise due to the specification error of parametric models, given their limited flexibility. Moreover, model misspecification is more likely in the presence of missing values. Different from the previous works where the informative variables are pre-specified, we propose a novel multi-task learning method to disentangle a learnable auxiliary representation, thereby ensuring the identifiability of the MNAR propensity.

Debiased Recommendations. Recommender systems are considered an effective tool for solving the information overload problem by searching structured data such as user features, item attributes, and user-item interactions stored in a database, and then selecting relevant items for each user [75]–[79]. However, the interactions recorded in the database were found to contain various biases [10], such as popularity bias [80], model selection bias [81], user self-selection bias [82], position bias [83], conformity bias [84], and confounding bias [42]. Those biases will result in the distribution of training data (observed population) different from that of test data (target population), leading to the challenges of achieving unbiased learning [85]–[92].

To achieve unbiased learning, many methods have been proposed. For example, [6] proposed the IPS method to debiasing for explicit feedback, and [93] extended it to the implicit feedback. [15] proposed a doubly robust joint learning (DR-JL) approach by enhancing the IPS method with an extra error imputation model and update the imputation model and prediction model alternately. In addition, [21], [60], [94] proposed to use information theory to mitigate the bias issue, [1], [95], [96] attempted to guarantee prediction performance by modeling users' behavior, and [97]–[99] adopted contrastive learning to train prediction model.

The above pioneering works have inspired a series of literature in recent years. We now focus on the recent methods that toward improving the limitations of DR-JL, such as MRDR [17] for reducing the variance, DR-MSE [53] for achieving better bias-variance trade-off, TDR [50] for achieving robustness to inaccurate pseudo-labelings, SDR [49] for reducing the bias, variance, and enhancing the robustness to small propensities, DR-V2 [61] for learning balanced propensities, N-DR [100] for addressing selection bias under neighborhood effect, CDR [101] for eliminating the poisonous imputations, CounterCLR [102] for aligning factual and counterfactual prediction, and AKBDR [103] for adaptively learning balanced propensities using kernel function. By combining multiple propensity models and imputation models, MR [52] proposed a multiple robust debiasing method that achieves unbiasedness if any of the propensity models or imputation models, or a linear combination of them is accurate. Besides, several multi-task learning methods are developed for debiasing and show competitive performance empirically, such as ESMM [19], Multi-DR [16], and ESCM^2-DR [18]. Moreover, [35], [51], [104]–[106] used an extra small uniform set for improving the debiasing performance. Specifically, [51] adaptively assigned propensity weights to biased training ratings, [35] leveraged the small uniform set to optimize the debiasing parameters, [104] used knowledge distillation for counterfactual recommendation, [105] proposed a unified multi-task debiasing approach, and [106] adaptively learned balance coefficients of biased samples. Unlike most existing debiasing methods that are biased on data MNAR, we propose a novel disentanglement approach for unbiased prediction using data MNAR without imposing strong assumptions.

VIII. Conclusion

In this paper, we formally summarized three different missing data mechanisms, i.e., MCAR, MAR, and MNAR, and found that previous propensity-based debiasing methods for addressing selection bias only achieve unbiasedness under the data MCAR and MAR. To achieve unbiased learning with data MNAR, we emphasize three key steps: (1) Setting the target MNAR propensities to depend on both observation indicators and true ratings; (2) Ensuring the identifiability of MNAR propensities by obtaining the auxiliary variables by disentanglement; (3) Employing MF as the base model to train the propensity model and the prediction model simultaneously. Extensive semi-synthetic and real-world experiments demonstrate the superiority of the proposed method in terms of debiasing performance and training efficiency. A limitation of our method is that its validity depends on whether the auxiliary variable can be successfully disentangled. However, it is empirically difficult to obtain precise auxiliary variables.

REFERENCES

[1] Z. Wang, X. Chen, R. Zhou, Q. Dai, Z. Dong, and J.-R. Wen, "Sequential recommendation with user causal behavior discovery," in *ICDE*, 2023.

[2] W. Liu, Y. Xi, J. Qin, X. Dai, R. Tang, S. Li, W. Zhang, and R. Zhang, "Personalized diversification for neural re-ranking in recommendation," in *ICDE*, 2023.

[3] Y. Zhang, Y. Liu, H. Xiong, Y. Liu, F. Yu, W. He, Y. Xu, L. Cui, and C. Miao, "Cross-domain disentangled learning for e-commerce live streaming recommendation," in *ICDE*, 2023.

[4] H. He, X. He, Y. Peng, Z. Shan, and X. Su, "Firzen: Firing strict cold-start items with frozen heterogeneous and homogeneous graphs for recommendation," in *ICDE*, 2024.

[5] B. M. Marlin and R. S. Zemel, "Collaborative prediction and ranking with non-random missing data," in *RecSys*, 2009.

[6] T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, and T. Joachims, "Recommendations as treatments: Debiasing learning and evaluation," in *ICML*, 2016.

[7] J. Li, Y. Moskovitch, and H. Jagadish, "Detection of groups with biased representation in ranking," in *ICDE*, 2023.

[8] J. Jin, H. Li, F. Feng, S. Ding, P. Wu, and X. He, "Fairly recommending with social attributes: A flexible and controllable optimization approach," in *NeurIPS*, 2024.

[9] P. Wu, H. Li, Y. Deng, W. Hu, Q. Dai, Z. Dong, J. Sun, R. Zhang, and X.-H. Zhou, "On the opportunity of causal learning in recommendation systems: Foundation, estimation, prediction and challenges," in *IJCAI*, 2022.

[10] J. Chen, H. Dong, X. Wang, F. Feng, M. Wang, and X. He, "Bias and debias in recommender system: A survey and future directions," *ACM Transactions on Information Systems*, vol. 41, no. 3, pp. 1–39, 2023.

[11] W. Wang, Y. Zhang, H. Li, P. Wu, F. Feng, and X. He, "Causal recommendation: Progresses and future directions," in *SIGIR*, 2023.

[12] H. Luo, F. Zhuang, R. Xie, H. Zhu, and D. Wang, "A survey on causal inference for recommendation," *arXiv preprint arXiv:2303.11666*, 2023.

[13] J. Huang, H. Oosterhuis, and M. de Rijke, "It is different when items are older: Debiasing recommendations when selection bias and user preferences are dynamic," in *WSDM*, 2022.

[14] H. Liu, D. Tang, J. Yang, X. Zhao, H. Liu, J. Tang, and Y. Cheng, "Rating distribution calibration for selection bias mitigation in recommendations," in *WWW*, 2022.

[15] X. Wang, R. Zhang, Y. Sun, and J. Qi, "Doubly robust joint learning for recommendation on data missing not at random," in *ICML*, 2019.

[16] W. Zhang, W. Bao, X.-Y. Liu, K. Yang, Q. Lin, H. Wen, and R. Ramezani, "Large-scale causal approaches to debiasing post-click conversion rate estimation with multi-task learning," in *WWW*, 2020.

[17] S. Guo, L. Zou, Y. Liu, W. Ye, S. Cheng, S. Wang, H. Chen, D. Yin, and Y. Chang, "Enhanced doubly robust learning for debiasing post-click conversion rate estimation," in *SIGIR*, 2021.

[18] H. Wang, T.-W. Chang, T. Liu, J. Huang, Z. Chen, C. Yu, R. Li, and W. Chu, "ESCM^2: Entire space counterfactual multi-task model for post-click conversion rate estimation," in *SIGIR*, 2022.

[19] X. Ma, L. Zhao, G. Huang, Z. Wang, Z. Hu, X. Zhu, and K. Gai, "Entire space multi-task model: An effective approach for estimating post-click conversion rate," in *SIGIR*, 2018.

[20] Y. Saito, H. Sakata, and K. Nakata, "Doubly robust prediction and evaluation methods improve uplift modeling for observational data," in *SDM*, 2019.

[21] D. Liu, P. Cheng, H. Zhu, Z. Dong, X. He, W. Pan, and Z. Ming, "Mitigating confounding bias in recommendation via information bottleneck," in *RecSys*, 2021.

[22] M. Zhang, J. Yuan, Y. He, W. Li, Z. Chen, and K. Kuang, "Map: Towards balanced generalization of iid and ood through model-agnostic adapters," in *ICCV*, 2023.

[23] M. Zhang, S. Huang, W. Li, and D. Wang, "Tree structure-aware few-shot image classification via hierarchical aggregation," in *ECCV*, 2022.

[24] S. Bai, M. Zhang, W. Zhou, S. Huang, Z. Luan, D. Wang, and B. Chen, "Prompt-based distribution alignment for unsupervised domain adaptation," in *AAAI*, 2024.

[25] M. Zhang, H. Li, F. Wu, and K. Kuang, "Metacoco: A new few-shot classification benchmark with spurious correlation," in *ICLR*, 2024.

[26] S. Bonner and F. Vasile, "Causal embeddings for recommendation," in *RecSys*, 2018.

[27] R. J. A. Little and D. B. Rubin, *Statistical Analysis with Missing Data*. John Wiley & Sons, 2019.

[28] A. A. Tsiatis, *Semiparametric Theory and Missing Data*. Springer, 2006.

[29] K. Mohan and J. Pearl, "Graphical models for processing missing data," *Journal of the American Statistical Association*, vol. 116, no. 534, pp. 1023–1037, 2021.

[30] J. D. Kang and J. L. Schafer, "Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data," *Statistical Science*, vol. 22, pp. 523–539, 2007.

[31] Z. Tan, "Comment: understanding OR, PS and DR," *Statistical Science*, vol. 22, pp. 560–568, 2007.

[32] S. R. Seaman and S. Vansteelandt, "Introduction to double robust methods for incomplete data," *Statistical Science*, vol. 33, pp. 184–197, 2018.

[33] H. Steck, "Training and testing of recommender systems on data missing not at random," in *KDD*, 2010.

[34] J. Shao and L. Wang, "Semiparametric inverse propensity weighting for nonignorable missing data," *Biometrika*, vol. 103, pp. 175–187, 2016.

[35] J. Chen, H. Dong, Y. Qiu, X. He, X. Xin, L. Chen, G. Lin, and K. Yang, "Autodebias: Learning to debias for recommendation," in *SIGIR*, 2021.

[36] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, no. 4, pp. 669–688, 1995.

[37] ——, *Causality*. Cambridge university press, 2009.

[38] G. Molenberghs, G. Fitzmaurice, M. G. Kenward, A. Tsiatis, and G. Verbeke, *Handbook of Missing Data Methodology*. CRC Press, 2014.

[39] W. Li, D. Yao, C. Gong, X. Chu, Q. Jing, X. Zhou, Y. Zhang, Y. Fan, and J. Bi, "Causaltad: Causal implicit generative model for debiased online trajectory anomaly detection," in *ICDE*, 2024.

[40] Y. Zhang, A. Wu, K. Kuang, L. Du, Z. Sun, and Z. Wang, "Stable heterogeneous treatment effect estimation across out-of-distribution populations," in *ICDE*, 2024.

[41] F. Zhu, M. Zhong, X. Yang, L. Li, Y. Lu, T. Zhang, J. Zhou, C. Chen, F. Wu, G. Liu *et al.*, "Dcmt: a direct entire-space causal multi-task framework for post-click conversion estimation," in *ICDE*, 2023.

[42] B. Youngmann, M. Cafarella, Y. Moskovitch, and B. Salimi, "On explaining confounding bias," in *ICDE*, 2023.

[43] F. Shen, K. Heravi, O. Gomez, S. Galhotra, A. Gilad, S. Roy, and B. Salimi, "Causal what-if and how-to analysis using hyper," in *ICDE*, 2023.

[44] P. K. Dan Siroker, *A/B Testing: The Most Powerful Way to Turn Clicks Into Customers*. John Wiley & Sons, 2015.

[45] A. Gilotte, C. Calauzènes, T. Nedelec, A. Abraham, and S. Dollé, "Offline a/b testing for recommender systems," in *WSDM*, 2018.

[46] C. A. Gomez-Uribe and N. Hunt, "The netflix recommender system: Algorithms, business value, and innovation," *ACM Transactions on Management Information Systems*, vol. 6, no. 4, pp. 1–19, 2015.

[47] R. Kohavi and R. Longbotham, "Online controlled experiments and a/b testing," *Encyclopedia of Machine Learning and Data Mining*, 2015.

[48] B. Yuan, J.-Y. Hsia, M.-Y. Yang, H. Zhu, C.-Y. Chang, Z. Dong, and C.-J. Lin, "Improving ad click prediction by considering non-displayed events," in *CIKM*, 2019.

[49] H. Li, C. Zheng, and P. Wu, "StableDR: Stabilized doubly robust learning for recommendation on data missing not at random," in *ICLR*, 2023.

[50] H. Li, Y. Lyu, C. Zheng, and P. Wu, "TDR-CL: Targeted doubly robust collaborative learning for debiased recommendations," in *ICLR*, 2023.

[51] X. Wang, R. Zhang, Y. Sun, and J. Qi, "Combating selection biases in recommender systems with a few unbiased ratings," in *WSDM*, 2021.

[52] H. Li, Q. Dai, Y. Li, Y. Lyu, Z. Dong, X.-H. Zhou, and P. Wu, "Multiple robust learning for recommendation," in *AAAI*, 2023.

[53] Q. Dai, H. Li, P. Wu, Z. Dong, X.-H. Zhou, R. Zhang, X. He, R. Zhang, and J. Sun, "A generalized doubly robust learning framework for debiasing post-click conversion rate prediction," in *KDD*, 2022.

[54] J. Qin, D. Leung, and J. Shao, "Estimation with survey data under nonignorable nonresponse or informative sampling," *Journal of the American Statistical Association*, vol. 97, no. 457, pp. 193–200, 2002.

[55] S. Wang, J. Shao, and J. K. Kim, "An instrumental variable approach for identification and estimation with nonignorable nonresponse," *Statistica Sinica*, pp. 1097–1116, 2014.

[56] Y. Wang, D. Liang, L. Charlin, and D. M. Blei, "Causal inference for recommender systems," in *RecSys*, 2020.

[57] Z. Zhu, X. Hu, and J. Caverlee, "Fairness-aware tensor-based recommendation," in *CIKM*, 2018.

[58] Y. Saito, "Doubly robust estimator for ranking metrics with post-click conversions," in *RecSys*, 2020.

[59] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[60] Z. Wang, X. Chen, R. Wen, S.-L. Huang, E. Kuruoglu, and Y. Zheng, "Information theoretic counterfactual learning from missing-not-at-random feedback," *NeurIPS*, 2020.

[61] H. Li, Y. Xiao, C. Zheng, P. Wu, and P. Cui, "Propensity matters: Measuring and enhancing balancing for recommendation," in *ICML*, 2023.

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

[63] W. Kweon and H. Yu, "Doubly calibrated estimator for recommendation on data missing not at random," in *WWW*, 2024.

[64] H. Li, Y. Liao, Z. Tian, Z. Liu, J. Liu, and X. Liu, "Bidirectional stackable recurrent generative adversarial imputation network for specific emitter missing data imputation," *IEEE Transactions on Information Forensics and Security*, 2024.

[65] G. Tang, R. J. Little, and T. E. Raghunathan, "Analysis of multivariate missing data with nonignorable nonresponse," *Biometrika*, vol. 90, no. 4, pp. 747–764, 2003.

[66] W. Miao, P. Ding, and Z. Geng, "Identifiability of normal and normal mixture models with nonignorable missing data," *Journal of the American Statistical Association*, vol. 111, no. 516, pp. 1673–1683, 2016.

[67] W. Miao, L. Liu, E. T. Tchetgen, and Z. Geng, "Identification, doubly robust estimation, and semiparametric efficiency theory of nonignorable missing data with a shadow variable," *arXiv preprint arXiv:1509.02556*, 2015.

[68] J. Zhao and J. Shao, "Semiparametric pseudo-likelihoods in generalized linear models with nonignorable missing data," *Journal of the American Statistical Association*, vol. 110, no. 512, pp. 1577–1590, 2015.

[69] E. J. Tchetgen Tchetgen and K. E. Wirth, "A general instrumental variable framework for regression analysis with outcome missing not at random," *Biometrics*, vol. 73, no. 4, pp. 1123–1131, 2017.

[70] B. Sun, L. Liu, W. Miao, K. Wirth, J. Robins, and E. J. T. Tchetgen, "Semiparametric estimation with data missing not at random using an instrumental variable," *Statistica Sinica*, vol. 28, no. 4, pp. 1965–1983, 2018.

[71] L. Liu, W. Miao, B. Sun, J. Robins, and E. Tchetgen Tchetgen, "Identification and inference for marginal average treatment effect on the treated with an instrumental variable," *Statistica Sinica*, vol. 30, no. 3, pp. 1517–1541, 2020.

[72] J. Zhao and Y. Ma, "A versatile estimation procedure without estimating the nonignorable missingness mechanism," *Journal of the American Statistical Association*, vol. 117, no. 540, pp. 1916–1930, 2022.

[73] J. S. Greenlees, W. S. Reece, and K. D. Zieschang, "Imputation of missing values when the probability of response depends on the variable being imputed," *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 251–261, 1982.

[74] J. K. Kim and C. L. Yu, "A semiparametric estimation of mean functionals with nonignorable missing data," *Journal of the American Statistical Association*, vol. 106, no. 493, pp. 157–165, 2011.

[75] Y. Zhang, C. Li, I. W. Tsang, H. Xu, L. Duan, H. Yin, W. Li, and J. Shao, "Diverse preference augmentation with multiple domains for cold-start recommendations," in *ICDE*, 2022.

[76] S. Dai, N. Shao, J. Zhu, X. Zhang, Z. Dong, J. Xu, Q. Dai, and J.-R. Wen, "Modeling user attention in music recommendation," in *ICDE*, 2024.

[77] J. Wu, J. Chen, J. Wu, W. Shi, J. Zhang, and X. Wang, "Bsl: Understanding and improving softmax loss for recommendation," in *ICDE*, 2024.

[78] Z. Lv, W. Zhang, S. Zhang, K. Kuang, F. Wang, Y. Wang, Z. Chen, T. Shen, H. Yang, B. C. Ooi *et al.*, "Duet: A tuning-free device-cloud collaborative parameters generation framework for efficient device model generalization," in *WWW*, 2023.

[79] Z. Lv, W. Zhang, Z. Chen, S. Zhang, and K. Kuang, "Intelligent model update strategy for sequential recommendation," in *WWW*, 2024.

[80] Y. Zhang, F. Feng, X. He, T. Wei, C. Song, G. Ling, and Y. Zhang, "Causal intervention for leveraging popularity bias in recommendation," in *SIGIR*, 2021.

[81] B. Yuan, J.-Y. Hsia, M.-Y. Yang, H. Zhu, C.-Y. Chang, Z. Dong, and C.-J. Lin, "Improving ad click prediction by considering non-displayed events," in *CIKM*, 2019.

[82] Y. Saito, "Asymmetric tri-training for debiasing missing-not-at-random explicit feedback," in *SIGIR*, 2020.

[83] Q. Ai, K. Bi, C. Luo, J. Guo, and W. B. Croft, "Unbiased learning to rank with unbiased propensity estimation," in *SIGIR*, 2018.

[84] Y. Liu, X. Cao, and Y. Yu, "Are you influenced by others when rating? improve rating prediction by conformity modeling," in *RecSys*, 2016.

[85] J. Zhou, X. Cao, W. Li, L. Bo, K. Zhang, C. Luo, and Q. Yu, "Hinet: Novel multi-scenario & multi-task learning with hierarchical information extraction," in *ICDE*, 2023.

[86] P. Arsomngern, C. Long, S. Suwajanakorn, and S. Nutanong, "Self-supervised deep metric learning for pointsets," in *ICDE*, 2021.

[87] H. Wang, K. Kuang, L. Lan, Z. Wang, W. Huang, F. Wu, and W. Yang, "Out-of-distribution generalization with causal feature separation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 4, pp. 1758–1772, 2024.

[88] H. Wang, K. Kuang, H. Chi, L. Yang, M. Geng, W. Huang, and W. Yang, "Treatment effect estimation with adjustment feature selection," in *KDD*, 2023.

[89] H. Wang, W. Yang, L. Yang, A. Wu, L. Xu, J. Ren, F. Wu, and K. Kuang, "Estimating individualized causal effect with confounded instruments," in *KDD*, 2022.

[90] H. Zou, H. Wang, R. Xu, B. Li, J. Pei, Y. J. Jian, and P. Cui, "Factual observation based heterogeneity learning for counterfactual prediction," in *CCLR*, 2023.

[91] H. Wang, J. Fan, Z. Chen, H. Li, W. Liu, T. Liu, Q. Dai, Y. Wang, Z. Dong, and R. Tang, "Optimal transport for treatment effect estimation," in *NeurIPS*, 2024.

[92] S. Huang, H. Li, Q. Li, C. Zheng, and L. Liu, "Pareto invariant representation learning for multimedia recommendation," in *ACM-MM*, 2023.

[93] Y. Saito, S. Yaginuma, Y. Nishino, H. Sakata, and K. Nakata, "Unbiased recommender learning from missing-not-at-random implicit feedback," in *WSDM*, 2020.

[94] J. Cao, J. Sheng, X. Cong, T. Liu, and B. Wang, "Cross-domain recommendation to cold-start users via variational information bottleneck," in *ICDE*, 2022.

[95] L. Xia, C. Huang, Y. Xu, P. Dai, M. Lu, and L. Bo, "Multi-behavior enhanced recommendation with cross-interaction collaborative relation modeling," in *ICDE*, 2021.

[96] C. Gao, X. He, D. Gan, X. Chen, F. Feng, Y. Li, T.-S. Chua, and D. Jin, "Neural multi-task recommendation from multi-behavior data," in *ICDE*, 2019.

[97] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "Bpr: Bayesian personalized ranking from implicit feedback," in *UAI*, 2009.

[98] Q. Wan, X. He, X. Wang, J. Wu, W. Guo, and R. Tang, "Cross pairwise ranking for unbiased item recommendation," in *WWW*, 2022.

[99] X. Xie, F. Sun, Z. Liu, S. Wu, J. Gao, J. Zhang, B. Ding, and B. Cui, "Contrastive learning for sequential recommendation," in *ICDE*, 2022.

[100] H. Li, C. Zheng, S. Ding, F. Feng, X. He, Z. Geng, and P. Wu, "Be aware of the neighborhood effect: Modeling selection bias under interference for recommendation," in *ICLR*, 2024.

[101] Z. Song, J. Chen, S. Zhou, Q. Shi, Y. Feng, C. Chen, and C. Wang, "CDR: Conservative doubly robust learning for debiased recommendation," in *CIKM*, 2023.

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

[103] H. Li, Y. Xiao, C. Zheng, P. Wu, X. Chen, Z. Geng, and P. Cui, "Debiased collaborative filtering with kernel-based causal balancing," in *ICLR*, 2024.

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

[105] H. Li, K. Wu, C. Zheng, Y. Xiao, H. Wang, F. Feng, X. He, Z. Geng, and P. Wu, "Removing hidden confounding in recommendation: a unified multi-task learning approach," in *NeurIPS*, 2023.

[106] H. Li, Y. Xiao, C. Zheng, and P. Wu, "Balancing unobserved confounding with a few unbiased ratings in debiased recommendations," in *WWW*, 2023.