LARGE-SCALE CAUSAL APPROACHES TO DEBIASING POST-CLICK CONVERSION RATE ESTIMATION WITH MULTI-TASK LEARNING

Wenhao Zhang[*][†]
University of California, Los Angeles
wenhaoz@ucla.edu

Wentian Bao[*]
Alibaba Group
wentian.bwt@alibaba-inc.com

Xiao-Yang Liu
Columbia University
xl2427@columbia.edu

Keping Yang
Alibaba Group
shaoyao@alibaba-inc.com

Quan Lin
Alibaba Group
tieyi.lq@alibaba-inc.com

Hong Wen
Alibaba Group
qinggan.wh@alibaba-inc.com

Ramin Ramezani
University of California, Los Angeles
raminr@ucla.edu

ABSTRACT

Post-click conversion rate (CVR) estimation is a critical task in e-commerce recommender systems. This task is deemed quite challenging under industrial setting with two major issues: 1) selection bias caused by user self-selection, and 2) data sparsity due to the rare click events. A successful conversion typically has the following sequential events: "exposure → click → conversion". Conventional CVR estimators are trained in the click space, but inference is done in the entire exposure space. They fail to account for the causes of the missing data and treat them as missing at random. Hence, their estimations are highly likely to deviate from the real values by large. In addition, the data sparsity issue can also handicap many industrial CVR estimators which usually have large parameter spaces.

In this paper, we propose two principled, efficient and highly effective CVR estimators for industrial CVR estimation, namely, **Multi-IPW** and **Multi-DR**. The proposed models approach the CVR estimation from a causal perspective and account for the causes of missing not at random. In addition, our methods are based on the multi-task learning framework and mitigate the data sparsity issue. Extensive experiments on industrial-level datasets show that our methods outperform the state-of-the-art CVR models.

*K*eywords Conversion rate estimation, causal inference, selection bias, multi-task learning, recommender systems.

1 Introduction

Selection bias is a widely-recognized issue in recommender systems [1, 2, 3]. For example, music stream services usually suggest genres that have positive user feedbacks (e.g., favorite, share, and buy, etc.), and selectively ignore the ones that are rarely exposed to users [4]. In this paper, we study the selection bias that exists in the post-click conversion rate (CVR) estimation.

Problem formulation

Post-click conversion rate (CVR) estimation is a critical task in e-commerce recommender systems [5, 6]. A typical e-commerce transaction has the following sequential events: "exposure → click → conversion" [2]. Post-click conversion

---

[*]Both authors contributed equally to this research. Authorship order determined by coin flip.

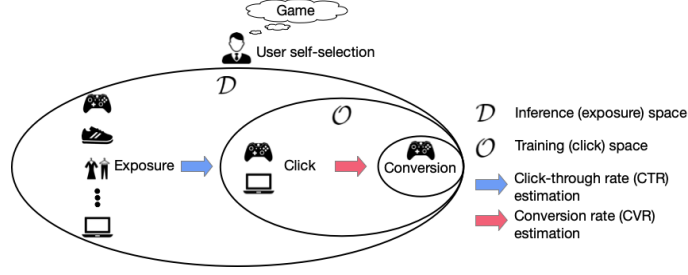[†]Work done while interning at Alibaba Group.

Figure 1: Illustration of the selection bias issue in conventional conversion rate (CVR) estimation. The training space of conventional CVR models is the click space $\mathcal{O}$, whereas the inference space is the entire exposure space $\mathcal{D}$. The discrepancy of data distribution between $\mathcal{O}$ and $\mathcal{D}$ leads to selection bias in conventional CVR models.

rate indicates the probability of transitions from click to conversion. Typically, when training CVR models, we only include the items that customers clicked on as we are unaware of the conversion feedback of the items that are not clicked by customers [7]. Bear in mind, not clicking on an item does not necessarily indicate the customer is not interested in purchasing it. Customers may unconsciously skip certain items that might be interesting to them. Fig. 1 reveals that the exposure space $\mathcal{D}$ is a super set of the click space $\mathcal{O}$. Selection bias occurs when conventional CVR models are trained in the click space, and the predictions are made in the entire exposure space (see Fig. 1) [2]. Intuitively, data in the click space is drawn from the entire exposure space and is biased by the user self-selection. Therefore, the data distribution in the click space is systematically different from the one in the exposure space. This inherent discrepancy leads to data that is missing not at random (MNAR), and selection bias in the conventional CVR models [3, 8, 9, 10].

In summary, we identify two practical issues that make CVR estimation quite challenging in industrial-level recommender systems:

- **Selection bias**: The systematic difference between data distributions in training space $\mathcal{O}$ and inference space $\mathcal{D}$ biases conventional CVR models [11, 7, 12, 10]. This bias usually causes the performance degradation.

- **Data sparsity**: This issue occurs since clicks are relatively rare events (we have a CTR of 5.2% in the production dataset and 4% in the public dataset). Conventional CVR models are typically trained only using data in the click space. Therefore, the number of training samples may not be sufficient for the large parameter space. In the experiments, we have 0.6 billion data samples vs. 5.3 billion parameters for production dataset, and 4.3 million samples vs. 2.6 billion parameters for public dataset (see Section 4.1) [13, 14].

To simplify the debiasing task of CVR estimation, we assume the exposure space is the entire item space we are interested in (see Fig. 1) [2]. Such a relaxation is also made based on the postulation that most items are exposed at least once. Table 1 shows that our dataset contains 81.5 million items and 11.5 billion exposures, i.e., each item is exposed, on average, about 150 times.

To address the critical issues of selection bias and data sparsity in the CVR estimation, we take a causal perspective and develop causal methods in a multi-task learning framework. We propose two principled, efficient and highly effective CVR estimators, namely, **Multi**-task **I**nverse **P**ropensity **W**eighting estimator (**Multi-IPW**) and **Multi**-task **D**oubly **R**obust estimator (**Multi-DR**). Our methods are designed for unbiased CVR estimation. They also account for the data sparsity issue.

The main contributions of this paper are summarized as follows:

- To the best of our knowledge, this is the first paper that combines IPW-based and DR-based methods with multi-task learning. From a causal perspective, we aim to tackle the well-recognized issues (i.e., selection bias and data sparsity) in CVR estimation in concert.

- We highlight that the state-of-the-art CVR model, ESMM [2], is biased. Different from existing works, our methods adjust for MNAR data, and deal with the selection bias in a principled way. Meanwhile, we give mathematical proofs that the proposed methods are theoretically unbiased. The empirical study shows our approaches outperform ESMM and several state-of-the-art causal models, and demonstrates the efficiency of our methods in real industrial setting.

2 Related works

In this section, we review several existing works that attempt to tackle the selection bias issue in recommender systems. Meanwhile, we summarize how our methods are different from prior works.

Ma *et al.* [2] proposed the Entire Space Multi-task Model (ESMM) to remedy selection bias and data sparsity issues in the conversion rate (CVR) estimation. ESMM is trained in the entire exposure space, and it formulates CVR task as two auxiliary tasks, i.e., click-through rate (CTR) and click-through & conversion rate (CTCVR) estimations. However, we argue that ESMM is biased. The details of our argument are presented in Section 3.2.

Causal inference offers a way to adapt for the data generation process when we attempt to restore the information from MNAR data [9]. Schnabel *et al.* [1] proposed an IPW-based estimator for training and evaluating recommender systems from biased data. IPW-based models may still be biased if the propensities are not accurately estimated. Wang *et al.* [15] proposed a doubly robust (DR) joint learning approach for estimating item ratings that are MNAR. Doubly robust estimator combines the IPW-based methods with an imputation model that estimates the prediction error for the missing data. When the propensities are not accurately learned, DR estimator can still enjoy unbiasedness as long as its imputation model is accurate. However, the existing DR-based methods are not devised for CVR estimation, hence fail to account for the severe data sparsity issue that widely exists in the CVR estimation. In addition, such a joint learning approach is not efficient in industrial setting (see Fig. 5).

To summarize, our approach differs from aforementioned methods in three aspects: 1) The problems are different. We developed our methods for CVR estimation in e-commerce system, while they focus on the rating prediction [16]. 2) The challenges are different. we design our models to address the selection bias and data sparsity issues, while they only consider the former (ESMM considers both). 3) The methods are different. we integrate multi-task framework with causal approaches. Specifically, We co-train propensity model, imputation model and prediction model simultaneously with deep neural networks, while they train these modules separately or alternatively, and usually with models such as linear regression or matrix factorization [17, 18, 19, 20]. We will further justify our design in Section 3 and report the performance improvement in Section 5.

3 Causal CVR Estimators with multi-task learning

3.1 Preliminary

Let $\mathcal{U} = (u_1, u_2, ..., u_N)$ be a set of N users and $\mathcal{I} = (i_1, i_2, ..., i_M)$ be a set of M items, $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ be the user-item pairs, $\mathbf{R} \in \mathbb{R}^{N \times M}$ be the true conversion label matrix where each entry $r_{u,i} \in \{0, 1\}$, and $\hat{\mathbf{R}} \in \mathbb{R}^{N \times M}$ be the predicted conversion score matrix where each entry $\hat{r}_{u,i} \in [0, 1]$. Then, the *Prediction inaccuracy* $\mathcal{P}$ over all user-item pairs can be formulated as follows,

$$\mathcal{P} = \mathcal{P}(\mathbf{R}, \hat{\mathbf{R}}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} e(r_{u,i}, \hat{r}_{u,i}), \tag{1}$$

where $e(r_{u,i}, \hat{r}_{u,i}) = -r_{u,i} \log(\hat{r}_{u,i}) - (1 - r_{u,i}) \log(1 - \hat{r}_{u,i})$.

Let $\mathbf{O} \in \{0, 1\}^{\mathcal{U} \times \mathcal{I}}$ be the *indicator matrix* where each entry $o_{u,i}$ is an observation indicator: $o_{u,i} = 1$ if a user u clicks on item i, $o_{u,i} = 0$ otherwise. Since the clicks are subjective to certain unobserved factors (e.g., users latent interests), such user self-selection process generates MNAR data [21, 22]. Naive CVR estimators are trained only in the click space $\mathcal{O} = \{(u, i) | o_{u,i} = 1, (u, i) \in \mathcal{D}\}$. Let $\mathbf{R}^{\text{obs}}$ and $\mathbf{R}^{\text{mis}}$ be the set of conversion labels that are present and absent in $\mathcal{D}$. We evaluate these naive CVR models by averaging the cross-entropy loss over the observed data [1, 15],

$$\begin{aligned}
\mathcal{E}^{\text{Naive}} &= \mathcal{E}(\mathbf{R}^{\text{obs}}, \hat{\mathbf{R}}) \\
&= \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{O}} e(r_{u,i}, \hat{r}_{u,i}) \\
&= \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{D}} o_{u,i} e(r_{u,i}, \hat{r}_{u,i}),
\end{aligned} \tag{2}$$

where $|\mathcal{O}| = \sum_{(u,i) \in \mathcal{D}} o_{u,i}$.

We say a CVR estimator $\mathcal{M}$ is *unbiased* when the expectation of the estimated prediction inaccuracy over $\mathbf{O}$ equals to the prediction inaccuracy $\mathcal{P}$, i.e., $\text{Bias}^{\mathcal{M}} = |\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{\mathcal{M}}] - \mathcal{P}| = 0$, otherwise it is *biased*. If data is MNAR, $|\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathcal{M}}] - \mathcal{P}| \gg 0$ [23].
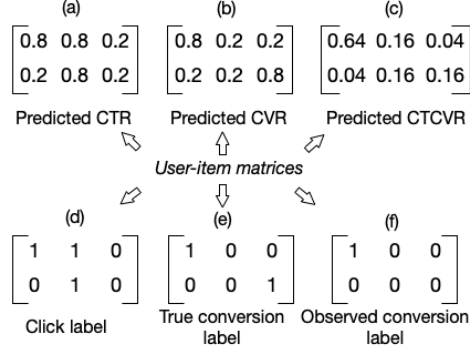
Figure 2: A toy example that demonstrates ESMM is biased.

3.2 Is ESMM an Unbiased CVR Estimator?

In this section, we demonstrate that ESMM, the state-of-the-art CVR estimator in practice, is essentially biased, though the author claim in the paper that the model eliminates the selection bias [2]. We formulate the estimation bias of ESMM, and prove it is not theoretically unbiased by giving a counter example.

Let $e_{u,i}^{\text{CTR}}, e_{u,i}^{\text{CVR}}, e_{u,i}^{\text{CTCVR}}, (u,i) \in \mathcal{D}$, be the cross-entropy losses of CTR, CVR, and CTCVR tasks. Then we have,

$$
\begin{aligned}
\text{Bias}^{\text{ESMM}} &= |\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{\text{ESMM}}] - \mathcal{P}| \\
&= \left| \mathbb{E}_{\mathbf{O}} \left[\frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} (e_{u,i}^{\text{CTR}} + e_{u,i}^{\text{CTCVR}}) \right] - \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e_{u,i}^{\text{CVR}} \right| \\
&= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} (e_{u,i}^{\text{CTR}} + e_{u,i}^{\text{CTCVR}}) - \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e_{u,i}^{\text{CVR}} \right| \\
&= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i)\in\mathcal{D}} (e_{u,i}^{\text{CTR}} + e_{u,i}^{\text{CTCVR}} - e_{u,i}^{\text{CVR}}) \right|.
\end{aligned}
\tag{3}
$$

We can easily verify that $\text{Bias}^{\text{ESMM}} > 0$ using the counter example in Fig. 2. Note that to be theoretically unbiased, ESMM should satisfy $|\mathbb{E}_{\mathcal{D}}[\mathcal{E}^{\text{ESMM}}] - \mathcal{P}| = 0, \forall \mathcal{D}$. Therefore, we conclude that ESMM cannot ensure unbiased CVR estimation.

3.3 A causal perspective to unbiased CVR estimation

Recall that selection bias in CVR estimation comes from the fact that models are trained over the click space $\mathcal{O}$, whereas the predictions are made over the exposure space $\mathcal{D}$ (See Fig. 1). Ideally, we can remove the selection bias by building our CVR estimators using a dataset where the conversion labels of all the items are known. In the language of causal inference, it is equivalent to training CVR estimators on a "do dataset", where causal intervention is applied on click event during the data generation process. Specifically, users are "forced" to click on every item in the exposure space $\mathcal{D}$ and further make their purchase decisions. Note that the training space is the same as the inference space in the "do dataset". Hence, the selection bias is eliminated. Intuitively, we can also understand how causal intervention removes the bias in Fig. 3. Z denotes the self-selection factors that affect both click events and conversion events. For example, Z can be the purchase interest or price discount that customers consider in online-shopping. In causal inference, we refer Z as "confounder(s)" that biases the CVR inference [24]. Once the causal intervention is applied on the click event (i.e., users are forced to click on all exposed items), Z has no control over user click behaviors. It means that we have successfully removed the confounder Z which biases our CVR estimation [25, 26, 24, 27, 28].

Apparently, this "do dataset" generated in this imaginary intervention experiment cannot be obtained in reality. Now the challenge is how to train our CVR estimators on the observed dataset $\mathcal{O}$ as if we do on the "do dataset". In the following sections, we will discuss two estimators that can achieve unbiased CVR prediction with the data that are MNAR.

As for ESMM and its variants, they are directly trained over the entire exposure space, which is different from the "do dataset". In "do dataset", we apply causal intervention on the click events during the data generation process to obtain the true conversion label matrix (see matrix e in Fig. 2). Without such intervention, the conversion labels in the dataset
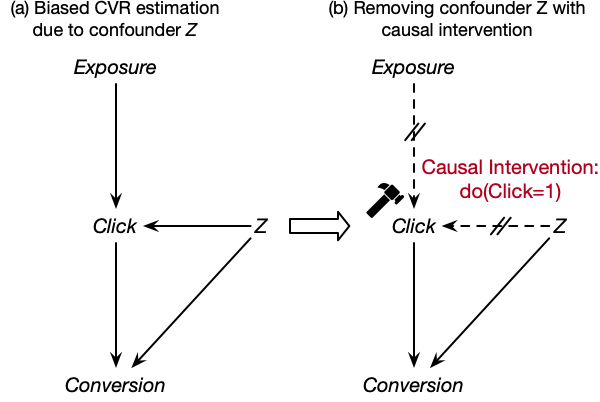
Figure 3: This causal graph formulate CVR estimation as a causal problem. [26] In (a), Z is a confounder that affects both clicks and purchases, and it biases the inference. In (b), we apply intervention on click events (do(Click)=1). Once users are "forced" to click on each exposed item, Z has no control over user click behaviors. Note the absence of the arrow from Z to *Click*. Hence, we have successfully removed the confounder Z, and the selection bias [25, 26, 24, 27, 28].

are still missing not at random. For example, in Fig. 2, matrix f is obtained without causal intervention. The conversion label of the last entry is observed as 0 because its click label is 0, whereas the true conversion label is 1 in the matrix e. We can also understand the weakness of ESMM and its variant from using causal graph (see Fig. 3), CVR estimators cannot achieve unbiased CVR estimation as long as they fail to remove the confounder Z.

3.4 Multi-task Learning Module

To address the data sparsity issue, we adopt the philosophy of multi-task learning and introduce an auxiliary CTR task [29]. The multi-task learning module exploits the typical sequential events in e-commerce recommender system, i.e., "exposure $\rightarrow$ click $\rightarrow$ conversion", and chains the main CVR task with the auxiliary CTR task. The amount of training data in CTR task is generally larger than that in CVR task by $1 \sim 2$ order of magnitudes (see Table 1), thus CTR task trains the large volume of model parameters more sufficiently. Besides, the feature representation learned in the CTR task is shared with the CVR task, which makes the CVR model benefit from the extra information via parameter sharing. Hence, the data sparsity issue is remedied [30, 2, 31].

Meanwhile, multi-task learning is also perceived as being cost-effective in training phase [2]. Specifically, multi-task learning co-trains multiple tasks simultaneously as if they were one task. This mechanism can potentially reduce storage space for saving duplicate copies of embedding matrix. In addition, the parallel training mechanism generally reduces the training time by large. The Multi-IPW and Multi-DR models inherit aforementioned merits by incorporating a multi-task learning module.

3.5 Multi-task Inverse Propensity Weighting CVR Estimator

Let the marginal probability $P(o_{u,i} = 1)$ denote the propensity score, $p_{u,i}$, of observing an entry in $\mathbf{R}$. In practice, the real $p_{u,i}$ can not be obtained directly. Instead, we estimate the real propensity with $\hat{p}_{u,i}$. The IPW-based estimator uses $\hat{p}_{u,i}$ to inversely weight prediction loss [9, 32, 1, 33],

$$\mathcal{E}^{\text{IPW}} = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} e(r_{u,i}, \hat{r}_{u,i})}{\hat{p}_{u,i}}. \tag{4}$$

Typically, $\hat{p}_{u,i}$ is learned via an independent logistic regression model [34]. In Fig. 4, we propose the Multi-IPW model which leverages the multi-task learning framework to simultaneously learn the propensity score (i.e., CTR in Multi-IPW) with CVR. Hence, the loss function of Multi-IPW estimator can be written as follows,

$$\mathcal{E}^{\text{Multi-IPW}}(\mathcal{X}_{\mathcal{O}}; \theta_{\text{CTR}}, \theta_{\text{CVR}}, \Phi)$$
$$= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \frac{o_{u,i} e(r_{u,i}, \hat{r}_{u,i}(\vec{x}_{u,i}; \theta_{\text{CVR}}, \Phi))}{\hat{p}_{u,i}(\vec{x}_{u,i}; \theta_{\text{CTR}}, \Phi)}, \tag{5}$$
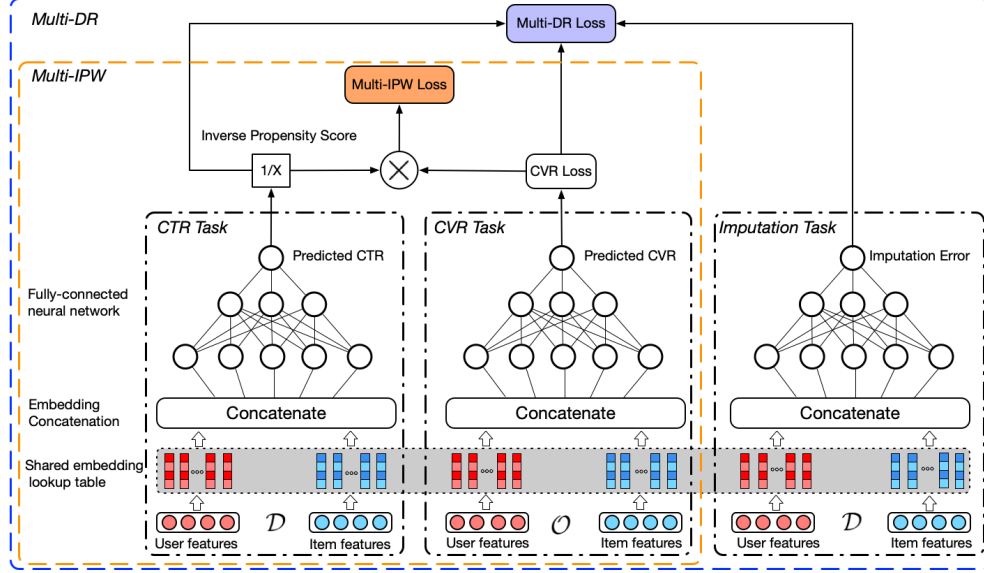
5

Figure 4: Multi-Inverse Propensity Weighting estimator and Multi-Doubly Robust estimator. The Multi-DR estimator augments Multi-IPW with an imputation model. We use predicted CTR as propensity scores in the Multi-IPW estimator. In the multi-task learning module, the CTR task, CVR task, and Imputation task are chained together via parameter sharing.

where Φ represents the shared embedding parameters. θ_{CVR} and θ_{CTR} are neural network parameters of CVR task and CTR task, respectively. $e(r_{u,i}, \hat{r}_{u,i})$, parameterized by θ_{CVR} and Φ, is the cross entropy loss of true CVR label $r_{u,i}$ and predicted CVR score $\hat{r}_{u,i}$. We use the predicted CTR score $\hat{p}_{u,i}$, parameterized by θ_{CTR} and Φ, as propensities. $\mathcal{D}$ denotes all the data in the exposure space. $\mathcal{X}_{\mathcal{O}}$ is the input feature vectors in $\mathcal{O}$.

we formally derive the bias of Multi-IPW and prove it is unbiased given the propensities are accurately estimated.

Theorem 1. *Given the true propensities* $\mathbf{P}$ *and the true conversion label matrix* $\mathbf{R}$, *the Multi-IPW CVR estimator gives unbiased CVR prediction when estimated propensity scores are accurate* $\hat{p}_{u,i} = p_{u,i}$,

$$|\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{Multi\text{-}IPW}] - \mathcal{P}| = 0. \tag{6}$$

Proof.

$$
\begin{aligned}
&|\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{\text{Multi-IPW}}] - \mathcal{P} \\
&= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \mathbb{E}_{\mathbf{O}}\left[\frac{o_{u,i}e(r_{u,i}, \hat{r}_{u,i})}{\hat{p}_{u,i}} \right] - \mathcal{P} \right| \\
&= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{p_{u,i}e(r_{u,i}, \hat{r}_{u,i})}{\hat{p}_{u,i}} - \mathcal{P} \right| \\
&= \left| \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e(r_{u,i}, \hat{r}_{u,i}) - \mathcal{P} \right| = 0.
\end{aligned} \tag{7}
$$

$\square$

Multi-IPW estimator inherits the merits of multi-task learning: 1) better CVR prediction due to parameter sharing, and 2) reduced training time and parameter storage. These are clear advantages over conventional IPW-based estimators.

3.6 Multi-task Doubly Robust CVR Estimator

The IPW-based models are unbiased contingent on accurately estimated propensities (i.e., $\hat{p}_{u,i} = p_{u,i}$). In practice, this condition is too restricted. To address this issue, doubly robust estimator is introduced by previous works [35, 15, 36, 37].

Algorithm 1: Multi-Inverse Propensity Weighting

Input: Observed conversion labels $\mathcal{R}^{obs}$ and user-item features $\mathcal{X}_{\mathcal{O}} = \{\vec{x}_{u,i}\}^{\mathcal{O}}, u, i \in \mathcal{O}$

while *stopping criteria is not satisfied* **do**

    Sample a batch $\mathcal{B}$ of user-item pairs $\{\vec{x}_{u,i}\}_B$ from $\mathcal{O}$

    Co-train CTR task and CVR task

        Update θ_{CTR}, Φ by descending along the gradients $\nabla_{\theta_{CTR}} \mathcal{E}^{Multi-IPW}, \nabla_{\Phi} \mathcal{E}^{Multi-IPW}$

        Update θ_{CVR}, Φ by descending along the gradients $\nabla_{\theta_{CVR}} \mathcal{E}^{Multi-IPW}, \nabla_{\Phi} \mathcal{E}^{Multi-IPW}$

end

Wang *et al.* [15] proposed a joint learning approach for training a doubly robust estimator, and introduced two models: 1) a prediction model $\hat{r}_{u,i} = f_\theta(\vec{x}_{u,i})$, and 2) an imputation model $\hat{e}_{u,i} = g_\phi(\vec{x}_{u,i})$. The prediction model, parameterized by θ, aims to predict the ratings, and its performance is evaluated by $e_{u,i} = e(r_{u,i}, \hat{r}_{u,i}), (u, i) \in \mathcal{D}$. The imputation model, parameterized by ϕ, aims to estimate the prediction error $e_{u,i}$ with $\hat{e}_{u,i}$. Its performance is assessed by $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i}$. The feature vector $\vec{x}_{u,i}$ encodes all the information about the user u and the item $i, (u, i) \in \mathcal{D}$. Then, we can formulate the loss of doubly robust estimator as,

$$\mathcal{E}^{\text{DR}} = \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \frac{o_{u,i} \delta_{u,i}}{\hat{p}_{u,i}} \right), \tag{8}$$

Similarly, we propose the Multi-DR estimator which augments Multi-IPW estimator by including an imputation model estimating the prediction error $e_{u,i}$. Multi-DR optimizes the following loss,

$$\mathcal{E}^{\text{Multi-DR}}(\mathcal{X}; \theta_{\text{CTR}}, \theta_{\text{CVR}}, \theta_{\text{Imp}}, \Phi)$$
$$= \frac{1}{|\mathcal{D}|} \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i}(\vec{x}_{u,i}; \theta_{\text{Imp}}, \Phi) + \frac{o_{u,i} \delta_{u,i}(\vec{x}_{u,i}; \theta_{\text{CVR}}, \theta_{\text{Imp}}, \Phi)}{\hat{p}_{u,i}(\vec{x}_{u,i}; \theta_{\text{CTR}}, \Phi)} \right), \tag{9}$$

where Φ represents the shared embedding parameters among CTR task, CVR task, and imputation task. $\theta_{\text{CTR}}, \theta_{\text{CVR}}, \theta_{\text{Imp}}$ are neural network parameters of CTR task, CVR task, imputation task, respectively. $\hat{p}_{u,i}$ is the propensity (i.e., predicted CTR score) given by CTR task. Estimated prediction error $\hat{e}$, parameterized by θ_{Imp} and Φ, is given by imputation task. $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i}$ is the error deviation.

We can formally derive the bias of Multi-DR and prove it is unbiased if either true propensity scores or true prediction errors are accurately estimated (i.e., $\Delta_{u,i} = 0$ or $\delta_{u,i} = 0$).

Theorem 2. *Given the true propensities* $\mathbf{P}$ *and the true conversion label matrix* $\mathbf{R}$*, the Multi-DR CVR estimator gives unbiased CVR prediction when either estimated propensity scores are accurate* $\Delta_{u,i} = \frac{p_{u,i} - \hat{p}_{u,i}}{\hat{p}_{u,i}} = 0$ *or the estimated prediction errors are accurate* $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i} = 0$,

$$|\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{\text{Multi-DR}}] - \mathcal{P}| = 0. \tag{10}$$

Proof.

$$|\mathbb{E}_{\mathbf{O}}[\mathcal{E}^{\text{Multi-DR}}] - \mathcal{P}|$$
$$= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \mathbb{E}_{\mathbf{O}} \left[\frac{o_{u,i} \delta_{u,i}}{\hat{p}_{u,i}} \right] \right) - \mathcal{P} \right|$$
$$= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \frac{p_{u,i} \delta_{u,i}}{\hat{p}_{u,i}} \right) - \mathcal{P} \right| \tag{11}$$
$$= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \frac{(p_{u,i} - \hat{p}_{u,i}) \delta_{u,i}}{\hat{p}_{u,i}} \right|$$
$$= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \Delta_{u,i} \delta_{u,i} \right| = 0.$$

$\square$

Algorithm 2: Multi-Doubly Robust

Input: Observed conversion labels $\mathcal{R}^{obs}$ and user-item features $\mathcal{X}_{\mathcal{D}} = \{\vec{x}_{u,i}\}^{\mathcal{D}}, u, i \in \mathcal{D}$
while *stopping criteria is not satisfied* **do**
    Sample a batch $\mathcal{B}$ of user-item pairs $\{\vec{x}_{u,i}\}_B$ from $\mathcal{D}$
    Co-train CTR task, CVR task, and Imputation task
        Update θ_{CTR}, Φ by descending along the gradients $\nabla_{\theta_{CTR}}\mathcal{E}^{Multi-DR}, \nabla_{\Phi}\mathcal{E}^{Multi-DR}$
        Update θ_{CVR}, Φ by descending along the gradients $\nabla_{\theta_{CVR}}\mathcal{E}^{Multi-DR}, \nabla_{\Phi}\mathcal{E}^{Multi-DR}$
        Update θ_{Imp}, Φ by descending along the gradients $\nabla_{\theta_{Imp}}\mathcal{E}^{Multi-DR}, \nabla_{\Phi}\mathcal{E}^{Multi-DR}$
end

---

4 Experimentation

In this section, we evaluate the performance of the proposed models with a public dataset and a large-scale production dataset collected from Mobile Taobao, the leading e-commerce platform in China. The experiments are intended to answer the following questions:

- **Q1**: Do our proposed approaches outperform other state-of-art CVR estimation methods?

- **Q2**: Are our proposed models more efficient in industrial setting than other baseline models?

- **Q3**: How is the performance of our proposed models affected by hyper-parameters?

4.1 Datasets

Ali-CCP [3][2]

Alibaba Click and Conversion Prediction (Ali-CCP) dataset is collected from real-world traffic logs of the recommender systems in the Taobao platform. See the statistics in Table 1.

Production sets

This industrial production dataset is collected from the Mobile Taobao e-commerce platform. It contains 3-week transactional data (see Table 1). Our production dataset includes 109 features, which are primarily categorized into: 1) user features, 2) item features, and 3) combination features. We further divide this dataset into 4 subsets: Set A, Set B, Set C, and Set D, which contain the first two days (5%), the first five days (20%), the first twelve days (50%), and the 3-week of data (100%), respectively. We use the data of the last day in each set as testing set and the remaining data as training set.

Table 1: Statistics of experimental datasets

Dataset	# Exposure	# Click	# Conversion	# user	# item
Ali-CCP	84M	3.4M	18k	0.4M	4.3M
Set A	1.1B	54.5M	0.6M	-	22.5M
Set B	2.7B	0.2B	1.9M	-	39.1M
Set C	6.0B	0.4B	4.3M	-	62.6M
Set D	11.5B	0.6B	8.3M	-	81.5M

4.2 Baseline models

We compare Multi-IPW model and Multi-DR model with the following baselines. Note that some baselines are causal estimators which we modify to predict the unbiased CVR, and others models are existing non-causal estimators designed for CVR predictions.

---

[3]https://tianchi.aliyun.com/dataset/dataDetail?dataId=408

4.2.1 Non-causal estimators

- **Base** is a naive post-click CVR model, which is a Multi-layer Perceptron (See the CVR task in Figure 4). Note that this is essentially an independent MLP model which takes the feature embeddings as input and predicts the CVR. The base model is trained in the click space.

- **Oversampling** [38, 39] deals with the class imbalance issue by duplicating the minority data samples (conversion=1) in training set with an oversampling rate k. In our experiment, we set k = 5. The oversampling model is trained in the click space.

- **ESMM** [2] utilizes multi-task learning methods and reduces the CVR estimation into two auxiliary tasks, i.e., CTR task and CTCVR task. ESMM is trained in the entire exposure space, and deemed as the state-of-the-art CVR estimation model in real practice.

- **Naive Imputation** takes all the unclicked data as negative samples. Hence, it is trained in the entire exposure space.

4.2.2 Causal estimators

- **Naive IPW**[1] is a naive IPW estimator. Note that it is not specifically designed for CVR estimation task as CVR prediction has its intrinsic issues. For example, it cannot deal with the data sparsity issue that inherently exists in CVR task.

- **Joint Learning DR** [15] is devised to learn from ratings that are missing not at random. In this experiment, we tailor Joint Learning DR for the CVR estimation. Similarly, Joint learning DR handles data sparsity issue poorly.

- **Heuristic DR** is designed as a baseline for Multi-DR. It assumes that the unclicked items are negative samples with probability $1 - \eta$, where η is smoothing rate and it denotes the probability of having a positive label. In the experiments, we explore η in $\{0.0005, 0.001, 0.002, 0.005, 0.01\}$ and report the best performance.

4.3 Metrics

In CVR prediction task, ROC AUC is a widely used metric [40]. One interpretation of AUC in the context of ranking system is that it denotes the probability of ranking a random positive sample higher than a negative sample. Meanwhile, we also adopt Group AUC (GAUC) [41]. GAUC extends AUC by calculating the weighted average of AUC grouped by page views or users,

$$\text{GAUC} = \frac{\sum_{i \in U} w_i \times \text{AUC}_i}{\sum_{i \in U} w_i}, \tag{12}$$

where w_i is exposures. GAUC is commonly recognized as a more indicative metric in real practice [41]. In the public dataset, models are only assessed with AUC as the dataset is missing the information for computing GAUC.

4.4 Unbiased Evaluation

In this work, we use CTCVR-AUC/GAUC to evaluate the unbiasedness of CVR estimators [2]. We need to point out that testing with an unbiased dataset or randomization is generally a golden standard for unbiasedness assessment [42, 43]. However, the unbiased training/testing dataset for CVR estimation is rather unobtainable in real practice.The real-world system can randomly expose items to users and generate unbiased evaluation sets for CTR estimation. But they cannot force users to randomly click on items to generate unbiased data for CVR estimations. This limitation may be further investigated in the future work.

4.5 Experiments setup

4.5.1 Ali-CCP experiment

The experiment setup on Ali-CCP mostly follows the prior work [2]. We set the dimension of all embedding vectors to be 18. The architecture of all these multi-layer perceptrons (MLP) in multi-task learning module are identical as $512 \times 256 \times 128 \times 32 \times 2$. The optimizer is Adam with a learning rate $lr = 0.0002$, and batch size is set to $|batch| = 1024$.

Table 2: Results of comparison study on Production datasets. The best scores are bold-faced in each column. Note that this table has two sections, AUC scores and GAUC scores. The rows that contain the models proposed in this paper are highlighted in color grey.

Model	Set A (1.1B)		Set B (2.7B)		Set C (6.0B)		Set D (11.5B)	
	CVR	CTCVR	CVR	CTCVR	CVR	CTCVR	CVR	CTCVR
AUC score								
Base	78.24	73.12	78.67	73.86	79.62	74.70	81.66	76.28
Oversampling[38]	78.63	73.53	78.72	74.09	79.69	74.82	81.77	76.30
ESMM[2]	79.29	73.86	79.74	74.33	80.11	74.97	82.17	76.55
Naive Imputation	78.12	73.21	78.44	73.50	79.32	73.81	81.56	76.39
Naive IPW[1]	79.23	73.82	79.73	74.34	80.14	74.92	82.13	76.45
Heuristic DR	78.45	73.45	78.84	73.99	79.52	74.18	81.74	76.40
Joint Learning DR[15]	79.09	73.67	79.53	74.51	80.01	74.90	82.09	76.61
Multi-IPW	79.51	73.99	**79.85**	74.81	80.21	75.01	82.57	76.89
Multi-DR	**79.72**	**74.45**	79.80	**74.91**	**80.50**	**75.39**	**82.72**	**77.23**
GAUC score								
Base	-	59.69	-	60.16	-	60.58	-	61.27
Oversampling[38]	-	60.17	-	60.28	-	60.59	-	61.30
ESMM[2]	-	60.53	-	60.90	-	61.13	-	61.76
Naive Imputation	-	60.14	-	60.39	-	60.56	-	61.39
Naive IPW[1]	-	60.51	-	60.95	-	61.09	-	61.77
Heuristic DR	-	60.01	-	60.30	-	60.65	-	61.35
Joint Learning DR[15]	-	60.43	-	60.83	-	60.97	-	61.67
Multi-IPW	-	60.70	-	**61.09**	-	61.25	-	61.98
Multi-DR	-	**60.90**	-	60.99	-	**61.52**	-	**62.28**

Table 3: Results of comparison study on Public dataset: Ali-CCP. Experiments are repeated 10 times and mean $\pm$ 1 std of AUC scores are reported below. The best scores are bold-faced in each column. The rows that contain the models proposed in this paper are highlighted in color grey.

Model	CVR AUC	CTCVR AUC
Base	66.00 ± 0.37	62.07 ± 0.45
Oversampling [38]	67.18 ± 0.32	63.05 ± 0.48
ESMM-NS [2]	68.25 ± 0.44	64.44 ± 0.62
ESMM [2]	68.56 ± 0.37	65.32 ± 0.49
Multi-IPW	69.21 ± 0.42	65.30 ± 0.50
Multi-DR	$\mathbf{69.29 \pm 0.31}$	$\mathbf{65.43 \pm 0.34}$

4.5.2 Production set experiment

In production set experiment, we vary the dimensions of feature embedding vectors according to each feature's real size in order to minimize the memory usage. In order to have a fair comparison study, all the models in this experiment share $|batch| = 10000$, MLP architecture $1024 \times 512 \times 256 \times 128 \times 32 \times 2$, adam optimizer with learning rate $lr = 0.0005$. We also added $l2$ normalization to imputation model in Multi-DR, and the coefficient is $v = 0.0001$.

5 Results and Discussion

In this section, we evaluate the proposed models and answer the questions raised in Section 4.

5.1 Model Assessments (Q1)

In this section, we report the experiment results in Table 2, 3. Multi-IPW and Multi-DR are clear winners over other baselines across all experiments. Meanwhile, we have the following observations:

- In production dataset, Multi-IPW and Multi-DR consistently outperform Joint Learning DR [15]. We reason that the performance improvement benefits from multi-task learning module, which remedies the data sparsity issue.

- We notice that Multi-DR mostly has better performance than Multi-IPW. Recall that Multi-DR augments Multi-IPW by introducing an imputation model. Provided that $0 \leq \hat{e} \leq 2e$, the tail bound of Multi-DR is proven to be lower than that of Multi-IPW for any learned propensity score $\hat{p}_{u,i}, (u,i) \in \mathcal{O}$ [15]. Therefore, Multi-DR is expected to perform better than Multi-IPW when the imputation model is well-trained.

- We observe that Multi-IPW/Naive IPW estimator are superior to Base in all experiments. Compared with the Base, both IPW-based models introduce estimated propensities to correct the selection bias. Recall that Theorem 1. ensures the CVR estimators are unbiased if the propensities are accurately estimated. While in practice the estimated propensities may deviate from the real values, this control experiment attests to "enough accuracy" of the propensity model.

The experiment results demonstrate that Multi-IPW and Multi-DR counter the selection bias and data sparsity issues in CVR estimation in a principled and highly effective way. In the next subsection, we will discuss other strengths of the proposed methods.
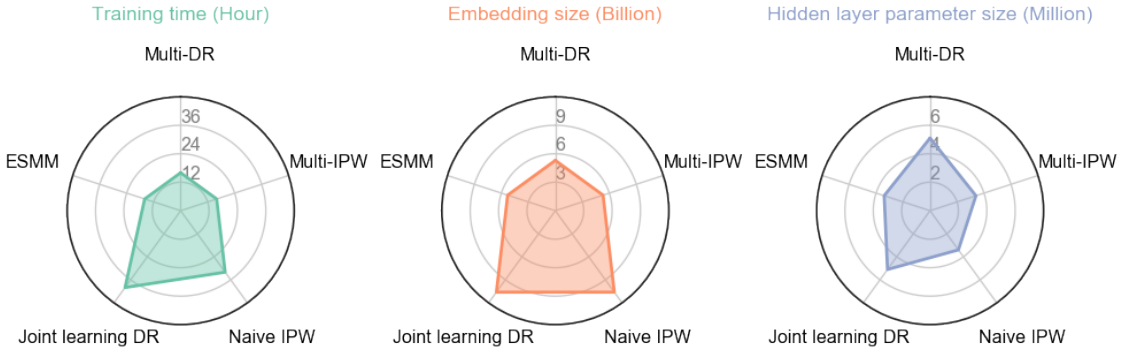
5.2 Computational efficiency (Q2)

Figure 5: Computational cost of Multi-IPW and Multi-DR. The left subplot reveals the hours needed to complete one epoch of training. The middle subplot shows the size of embedding parameters of each model. The right subplot shows the size of hidden layer parameters of each model. Note that the proposed models achieve the best prediction performance, while have the lowest computational cost.

In this section, we study the computational efficiency of the proposed models against the baselines under industrial setting. We summarize the records of training time and parameter space size of each model in Fig. 5, and the cluster configuration in Table 4.

We observe that Multi-IPW and Multi-DR require less or equivalent training time compared with other baselines. Recall that multi-task learning method co-trains multiple tasks simultaneously as if they were one task. We can expect the training time being greatly shortened. Meanwhile, our methods are also economical in memory usage due to parameter sharing in the multi-task learning module.

Table 4: Distributed cluster configuration

Cluster configuration	Parameter Server	Worker
# instances	4	100
# CPU	28 cores	440 cores
# GPU[4]	-	25 cards
MEMORY (GB)	40	1000

5.3 Hyper-parameters in model implementation

IPW bound τ is a hyper-parameter introduced in our model implementation to handle high variance of propensities. IPW bound clamps the propensities if the values are greater than the predefined threshold. A plausible IPW bound value is typically confined by $\tau \in [propensity_{min}, propensity_{mean}]$. IPW bound percentage can be calculated as

---

[4]GPU specs: Tesla P100-PCIE-16G

$\tau\% = \frac{\tau - propensity_{min}}{propensity_{mean} - propensity_{min}}$. In Multi-DR, imputation model will <u>introduce the unclicked items to the training set</u>. <mark>Empirically, most of the unclicked items will not be purchased by customers even if they were clicked</mark>. Therefore, including these unclicked items in training set will <u>skew the data distribution</u> and make the <u>class imbalance issue worse</u>. Therefore, Instead of adding all the unclicked samples, we <u>under-sample</u> them with a <u>sampling rate λ</u>. For example, if the number of clicked samples ($N_{clicked}$) is 100 and the batch size is 1000, $\lambda = 1.5$ means that after under-sampling the samples we used to train Multi-DR is $N_{clicked} \times \lambda = 100 \times 1.5 = 150$. Note that without under-sampling, Multi-DR takes all samples in the batch as training samples.

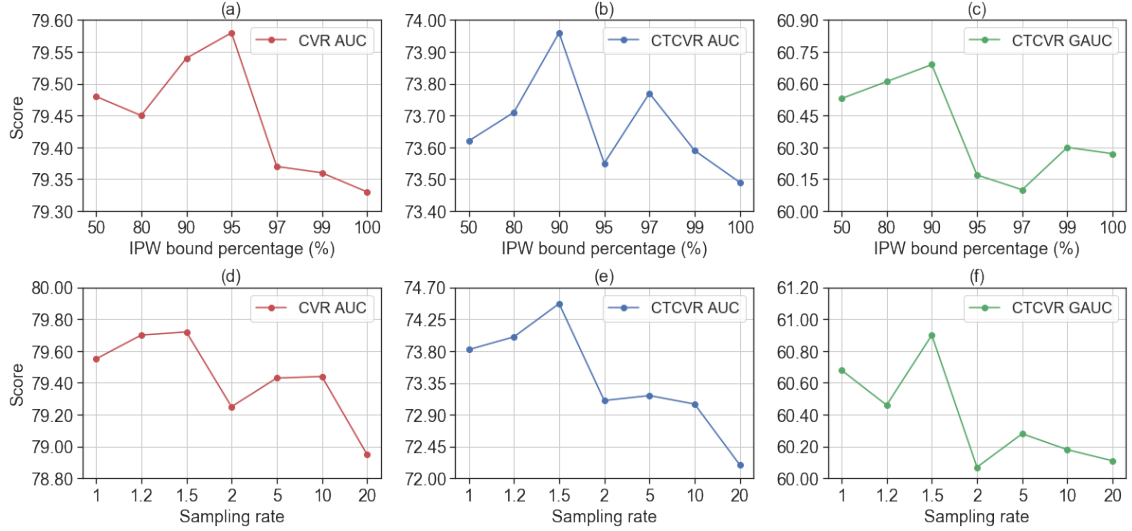## 5.4 Empirical study on hyper-Parameter sensitivity (Q3)

Figure 6: Results of parameter sensitivity experiments.

In this section, we investigate how Multi-IPW and Multi-DR are affected by two important hyper-parameters in our model implementation, IPW bound τ and sampling rate λ. We evaluate the performance of Multi-IPW with varying IPW bound τ. We observe that in Fig. 6, when IPW bound $50\% < \tau\% < 90\%$, prediction performance eventually improves as IPW bound increases. We can clearly see the performance drop of CVR AUC if the threshold is greater than 95%. We reason that <u>larger IPW bound allows undesired higher variability of propensity scores</u>, which may lead to sub-optimal prediction performance.

We evaluate the performance of Multi-DR with varying sampling rate λ. We observe that <u>$\lambda = 1.5$ produces the best prediction</u>, and the model performance starts decreasing when $\lambda > 5$. We argue that, as the sampling rate increases, more unclicked samples are included to our training set, and it <u>inevitably worsens the class imbalance issue</u>, which typically causes predictive models to generalize poorly. On the contrary, <u>introducing a small number of unclicked samples from the imputation model can boost our CVR prediction</u> (see figure (d) when $\lambda \in [1.0, 1.5]$).

6 Conclusion and future works

In this paper, we proposed <u>Multi-IPW</u> and <u>Multi-DR</u> CVR estimators for industrial recommender system. Both CVR estimators aim to counter the <u>inherent issues</u> in practice: 1) <u>selection bias</u>, and 2) <u>data sparsity</u>. Extensive experiments with billions of data samples demonstrate that our methods outperform the state-of-the-art CVR predictive models, and handle CVR estimation task in a principled, highly effective and efficient way. Although our methods are devised for CVR estimation, <u>the idea can be generalized to debiasing CTR estimation</u> by exploiting the sequential pattern "item pool $\rightarrow$ exposure $\rightarrow$ click".

References

[1] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. Recommendations as treatments: Debiasing learning and evaluation. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, ICML'16, page 1670–1679. JMLR.org, 2016.

[2] Xiao Ma, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. Entire space multi-task model: An effective approach for estimating post-click conversion rate. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '18, page 1137–1140, New York, NY, USA, 2018. Association for Computing Machinery.

[3] Arnaud De Myttenaere, Bénédicte Le Grand, Boris Golden, and Fabrice Rossi. Reducing offline evaluation bias in recommendation systems. In *23rd annual Belgian-Dutch Conference on Machine Learning (Benelearn 2014)*, pages 55–62, Bruxelles, Belgium, June 2014.

[4] Aaron Van den Oord, Sander Dieleman, and Benjamin Schrauwen. Deep content-based music recommendation. In *Advances in Neural Information Processing Systems*, pages 2643–2651, 2013.

[5] Han Zhu, Junqi Jin, Chang Tan, Fei Pan, Yifan Zeng, Han Li, and Kun Gai. Optimized cost per click in taobao display advertising. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2191–2200, 2017.

[6] Yuriko Yamaguchi, Mimpei Morishita, Youichi Inagaki, Reyn Nakamoto, Jianwei Zhang, Junichi Aoi, and Shinsuke Nakajima. Web advertising recommender system based on estimating users' latent interests. In *Proceedings of the 18th International Conference on Information Integration and Web-based Applications and Services*, pages 42–49, 2016.

[7] Jiayuan Huang, Arthur Gretton, Karsten Borgwardt, Bernhard Schölkopf, and Alex J Smola. Correcting sample selection bias by unlabeled data. In *Advances in Neural Information Processing Systems*, pages 601–608, 2007.

[8] Roderick JA Little and Donald B Rubin. *Statistical analysis with missing data*, volume 793. John Wiley & Sons, 2019.

[9] RJA Little and DB Rubin. Statistical analysis with missing data. wiley. *New York*, 2002.

[10] Elias Bareinboim, Jin Tian, and Judea Pearl. Recovering from selection bias in causal and statistical inference. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.

[11] Harald Steck. 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*, pages 713–722. ACM, 2010.

[12] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W Bruce Croft. Unbiased learning to rank with unbiased propensity estimation. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 385–394. ACM, 2018.

[13] Kuang-chih Lee, Burkay Orten, Ali Dasdan, and Wentong Li. Estimating conversion rate in display advertising from past erformance data. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 768–776. ACM, 2012.

[14] Jizhe Wang, Pipei Huang, Huan Zhao, Zhibo Zhang, Binqiang Zhao, and Dik Lun Lee. Billion-scale commodity embedding for e-commerce recommendation in alibaba. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 839–848. ACM, 2018.

[15] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. Doubly robust joint learning for recommendation on data missing not at random. *Proceedings of Machine Learning Research*, 2019.

[16] Yichao Lu, Ruihai Dong, and Barry Smyth. Coevolutionary recommendation model: Mutual learning between ratings and reviews. In *Proceedings of the 2018 World Wide Web Conference*, pages 773–782. International World Wide Web Conferences Steering Committee, 2018.

[17] Prem K Gopalan, Laurent Charlin, and David Blei. Content-based recommendations with poisson factorization. In *Advances in Neural Information Processing Systems*, pages 3176–3184, 2014.

[18] Huifeng Guo, Ruiming Tang, Yunming Ye, Zhenguo Li, and Xiuqiang He. Deepfm: a factorization-machine based neural network for ctr prediction. *arXiv preprint arXiv:1703.04247*, 2017.

[19] Steffen Rendle. Factorization machines. In *2010 IEEE International Conference on Data Mining*, pages 995–1000. IEEE, 2010.

[20] Xiangnan He, Hanwang Zhang, Min-Yen Kan, and Tat-Seng Chua. Fast matrix factorization for online recommendation with implicit feedback. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 549–558. ACM, 2016.

[21] Donald B Rubin. Inference and missing data. *Biometrika*, 63(3):581–592, 1976.

[22] Craig K Enders. *Applied missing data analysis*. Guilford press, 2010.

[23] Karthika Mohan, Judea Pearl, and Jin Tian. Graphical models for inference with missing data. In *Advances in Neural Information Processing Systems*, pages 1277–1285, 2013.

[24] Yixin Wang, Dawen Liang, Laurent Charlin, and David M Blei. The deconfounded recommender: A causal inference approach to recommendation. *arXiv preprint arXiv:1808.06581*, 2018.

[25] Judea Pearl et al. Causal inference in statistics: An overview. *Statistics surveys*, 3:96–146, 2009.

[26] J. Pearl and D. Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, 2018.

[27] Judea Pearl. *Causality: models, reasoning and inference*, volume 29. Springer, 2000.

[28] Stephen L Morgan and Christopher Winship. *Counterfactuals and causal inference*. Cambridge University Press, 2015.

[29] Hongwei Wang, Fuzheng Zhang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. Multi-task feature learning for knowledge graph enhanced recommendation. In *The World Wide Web Conference*, pages 2000–2010. ACM, 2019.

[30] Guy Hadash, Oren Sar Shalom, and Rita Osadchy. Rank and rate: multi-task learning for recommender systems. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 451–454. ACM, 2018.

[31] Yabo Ni, Dan Ou, Shichen Liu, Xiang Li, Wenwu Ou, Anxiang Zeng, and Luo Si. Perceive your users in depth: Learning universal user representations from multiple e-commerce tasks. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 596–605. ACM, 2018.

[32] Guido W Imbens and Donald B Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

[33] Keisuke Hirano, Guido W Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

[34] Peter C Austin. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3):399–424, 2011.

[35] Miroslav Dudík, John Langford, and Lihong Li. Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML'11, page 1097–1104, Madison, WI, USA, 2011. Omnipress.

[36] Karel Vermeulen and Stijn Vansteelandt. Bias-reduced doubly robust estimation. *Journal of the American Statistical Association*, 110(511):1024–1036, 2015.

[37] Mehrdad Farajtabar, Yinlam Chow, and Mohammad Ghavamzadeh. More robust doubly robust off-policy evaluation. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 1447–1456, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR.

[38] Gary M Weiss. Mining with rarity: a unifying framework. *ACM Sigkdd Explorations Newsletter*, 6(1):7–19, 2004.

[39] Wenhao Zhang, Ramin Ramezani, and Arash Naeim. WOTBoost: Weighted oversampling technique in boosting for imbalanced learning. *arXiv e-prints*, page arXiv:1910.07892, Oct 2019.

[40] Tom Fawcett. An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.

[41] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018.

[42] Ronald A Fisher. *Statistical methods and scientific inference.* Hafner Publishing Co., 1956.

[43] Anne Whitehead and John Whitehead. A general parametric approach to the meta-analysis of randomized clinical trials. *Statistics in Medicine*, 10(11):1665–1677, 1991.