



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 \rightarrow click \rightarrow 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.

CCS CONCEPTS

• **Information systems** Recommender systems; Information retrieval diversity; Personalization.

KEYWORDS

conversion rate estimation; causal inference; selection bias; multi-task learning; recommender systems

^{*}Both authors contributed equally. Authorship order determined by coin flip.

[†]Work done while interning at Alibaba Group.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '20, April 20–24, 2020, Taipei, Taiwan

© 2020 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-7023-3/20/04.

<https://doi.org/10.1145/3366423.3380037>

ACM Reference Format:

Wenhao Zhang, Wentian Bao, Xiao-Yang Liu, Keping Yang, Quan Lin, Hong Wen, and Ramin Ramezani. 2020. Large-scale Causal Approaches to Debiasing Post-click Conversion Rate Estimation with Multi-task Learning. In *Proceedings of The Web Conference 2020 (WWW '20)*, April 20–24, 2020, Taipei, Taiwan. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3366423.3380037>

1 INTRODUCTION

Selection bias is a widely-recognized issue in recommender systems [4, 17, 21]. 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 [23]. In this paper, we study the selection bias that exists in the post-click conversion rate (CVR) estimation.

Problem Statement. Post-click conversion rate (CVR) estimation is a critical task in e-commerce recommender systems [30, 33]. A typical e-commerce transaction has the following sequential events: "exposure \rightarrow click \rightarrow conversion" [17]. Post-click conversion 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 [12]. 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) [17]. 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, 4, 15, 16].

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

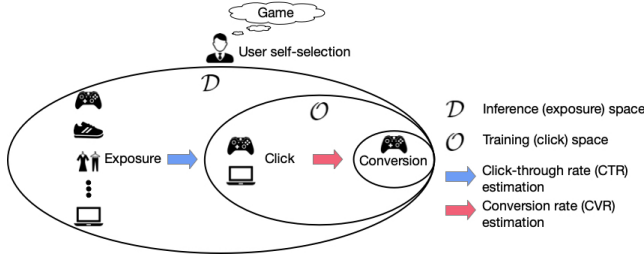


Figure 1: Illustration of the selection bias issue in conventional conversion rate (CVR) estimation. The training space of conventional CVR models is the click space O , whereas the inference space is the entire exposure space D . The discrepancy of data distribution between O and D leads to selection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

conventional CVR models [1, 3, 12, 22]. 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) [14, 26].

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) [17]. 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 Inverse Propensity Weighting estimator (Multi-IPW)** and **Multi-task Doubly Robust 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 [17], 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

Ma *et al.* [17] 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 [15]. Schnabel *et al.* [21] 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.* [27] 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. 4).

3 CAUSAL CVR ESTIMATORS WITH MULTI-TASK LEARNING

3.1 Preliminary

Let $\mathcal{U} = (u_1, u_2, \dots, u_N)$ be a set of N users and $\mathcal{I} = (i_1, i_2, \dots, i_M)$ 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,

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

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 [6, 20]. 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}^{obs} and \mathbf{R}^{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 [21, 27],

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

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

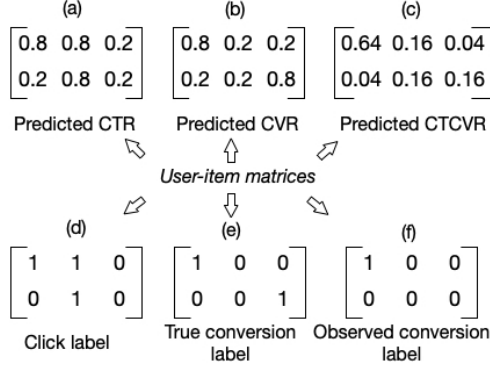


Figure 2: A toy example that demonstrates ESMM is biased.

We say a CVR estimator \mathcal{M} is *unbiased* when the expectation of the estimated prediction inaccuracy over \mathcal{O} equals to the prediction inaccuracy \mathcal{P} , i.e., $\text{Bias}^{\mathcal{M}} = |\mathbb{E}_{\mathcal{O}}[\mathcal{E}^{\mathcal{M}}] - \mathcal{P}| = 0$, otherwise it is *biased*. If data is MNAR, $|\mathbb{E}_{\mathcal{O}}[\mathcal{E}_{\mathcal{M}}] - \mathcal{P}| \gg 0$ [18].

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 [17]. 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,

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

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 Multi-task Learning Module

To address the data sparsity issue, we adopt the philosophy of multi-task learning and introduce an auxiliary CTR task [25]. 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 ~ 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 [10, 17, 19].

Meanwhile, multi-task learning is also perceived as being cost-effective in training phase [17]. 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.

3.4 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 [11, 13, 15, 21],

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

Typically, $\hat{p}_{u,i}$ is learned via an independent logistic regression model [2]. In Fig. 3, 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,

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

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 3.1. *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 $\hat{p}_{u,i} = p_{u,i}$,*

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

PROOF.

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

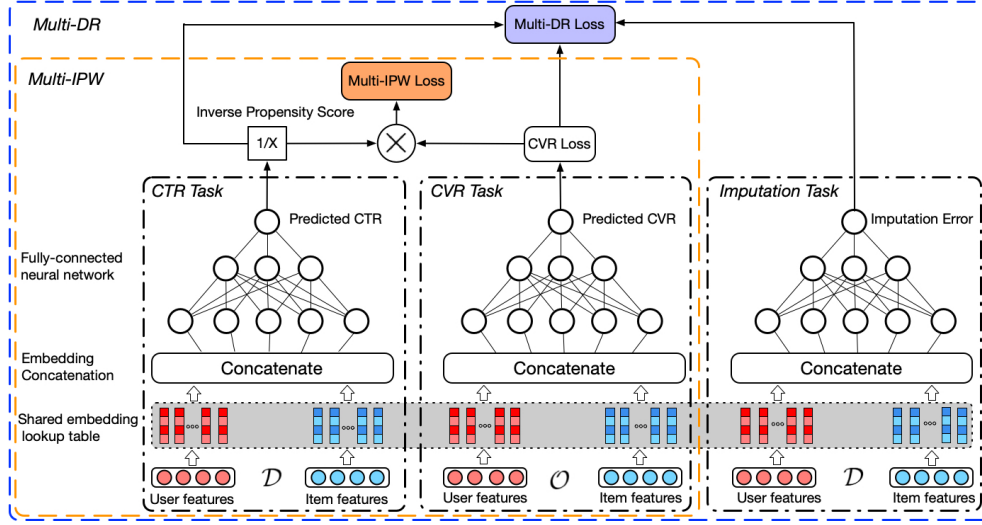


Figure 3: Multi-Inverse Propensity Weighting estimator and Multi-Doubly Robust estimator. The Multi-DR estimator augments Multi-IPW with an imputation model. We use predicted CTR as propensity scores in the Multi-IPW estimator. In the multi-task learning module, the CTR task, CVR task, and Imputation task are chained together via parameter sharing.

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.

3.5 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 [5, 7, 24, 27].

Wang *et al.* [27] 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,

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

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,

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

where Φ represents the shared embedding parameters among CTR task, CVR task, and imputation task. θ_{CTR} , θ_{CVR} , θ_{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 3.2. *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 $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i} = 0$,*

$$|\mathbb{E}_0[\mathcal{E}^{\text{Multi-DR}}] - \mathcal{P}| = 0. \quad (10)$$

PROOF.

$$\begin{aligned} & |\mathbb{E}_0[\mathcal{E}^{\text{Multi-DR}}] - \mathcal{P}| \\ &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \mathbb{E}_0 \left[\frac{o_{u,i} \delta_{u,i}}{\hat{p}_{u,i}} \right] \right) - \mathcal{P} \right| \\ &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \left(\hat{e}_{u,i} + \frac{p_{u,i} \delta_{u,i}}{\hat{p}_{u,i}} \right) - \mathcal{P} \right| \\ &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \frac{(p_{u,i} - \hat{p}_{u,i}) \delta_{u,i}}{\hat{p}_{u,i}} \right| \\ &= \frac{1}{|\mathcal{D}|} \left| \sum_{(u,i) \in \mathcal{D}} \Delta_{u,i} \delta_{u,i} \right| = 0. \end{aligned} \quad (11)$$

□

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 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 the state-of-the-art CVR estimation methods?
- **Q2:** Are our proposed models more efficient in industrial setting than other baseline models?

4.1 Datasets

*Ali-CCP*¹ [17]. 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.

4.2 Baseline Models

We compare the proposed models with the following baselines.

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 Fig. 3).
- **Oversampling** [28, 31] Duplicates the minority data samples (conversion=1) in training set with an oversampling rate $k=5$.
- **ESMM** [17] learns CVR using a CTR task and a CTCVR task.
- **Naive Imputation** uses unclicked data as negative samples.

4.2.2 Causal estimators.

- **Naive IPW** [21] is the naive IPW estimator (see Section 3.4).
- **Joint Learning DR** [27] is tailored for CVR estimation.
- **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 $\{0.0005, 0.001, 0.002, 0.005, 0.01\}$.

¹<https://tianchi.aliyun.com/dataset/dataDetail?dataId=408>

4.3 Metrics

In CVR prediction task, ROC AUC is a widely used metric [8]. 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) [32]. GAUC extends AUC by calculating the weighted average of AUC grouped by page views or users,

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

where w_i is exposures. GAUC is commonly recognized as a more indicative metric in real practice [32]. 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 [17]. We need to point out that testing with an unbiased dataset or randomization is generally a golden standard for unbiasedness assessment [9, 29]. 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

Ali-CCP experiment. The experiment setup on Ali-CCP follows the prior work [17].

Production set experiment. In the 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 $l2$ regularization to imputation model in Multi-DR, and the coefficient is $v = 0.0001$.

5 RESULTS AND DISCUSSION

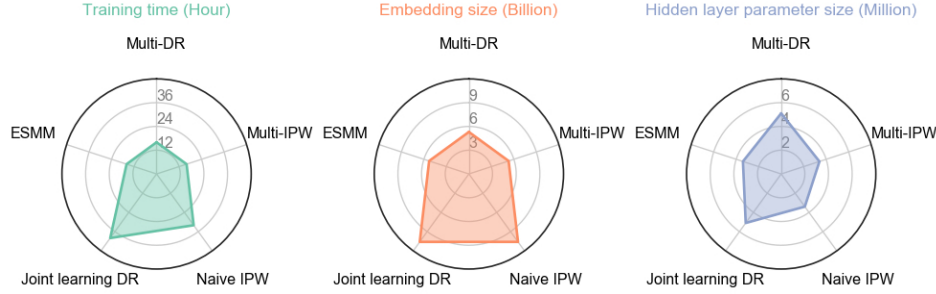
5.1 Model Assessments (Q1)

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 [27]. 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 $\hat{p}_{u,i}, (u, i) \in \mathcal{O}$ [27]. Therefore, Multi-DR is expected to perform better than Multi-IPW when the imputation model is well-trained.

Table 2: Results of comparison study on Production datasets. The best scores are bold-faced in each column.

Model	Set A (1.1B)			Set B (2.7B)			Set C (6.0B)			Set D (11.5B)		
	CVR AUC	CTCVR AUC	CTCVR GAUC	CTR AUC	CTCVR AUC	CTCVR GAUC	CTR AUC	CTCVR AUC	CTCVR GAUC	CTR AUC	CTCVR AUC	CTCVR GAUC
Base	78.24	73.12	59.69	78.67	73.86	60.16	79.62	74.70	60.58	81.66	76.28	61.27
Oversampling [28]	78.63	73.53	60.17	78.72	74.09	60.28	79.69	74.82	60.59	81.77	76.30	61.30
ESMM [17]	79.29	73.86	60.53	79.74	74.33	60.90	80.11	74.97	61.13	82.17	76.55	61.76
Naive Imputation	78.12	73.21	60.14	78.44	73.50	60.39	79.32	73.81	60.56	81.56	76.39	61.39
Naive IPW [21]	79.23	73.82	60.51	79.73	74.34	60.95	80.14	74.92	61.09	82.13	76.45	61.77
Heuristic DR	78.45	73.45	60.01	78.84	73.99	60.30	79.52	74.18	60.65	81.74	76.40	61.35
Joint Learning [27]	79.09	73.67	60.43	79.53	74.51	60.83	80.01	74.90	60.97	82.09	76.61	61.67
Multi-IPW	79.51	73.99	60.70	79.85	74.81	61.09	80.21	75.01	61.25	82.57	76.89	61.98
Multi-DR	79.72	74.45	60.90	79.80	74.91	60.99	80.50	75.39	61.52	82.72	77.23	62.28

**Figure 4: Computational cost comparison radar diagrams. The left subplot reveals the hours needed to complete one epoch of training. The middle subplot shows the size of embedding parameters. The right subplot shows the size of hidden layer parameters.****Table 3: Results of comparison study on Ali-CCP dataset. Experiments are repeated 10 times and mean \pm 1 std of AUC scores are reported. The best scores are bold-faced.**

Model	CVR AUC	CTCVR AUC
Base	66.00 \pm 0.37	62.07 \pm 0.45
Oversampling [28]	67.18 \pm 0.32	63.05 \pm 0.48
ESMM-NS [17]	68.25 \pm 0.44	64.44 \pm 0.62
ESMM [17]	68.56 \pm 0.37	65.32 \pm 0.49
Multi-IPW	69.21 \pm 0.42	65.30 \pm 0.50
Multi-DR	69.29 \pm 0.31	65.43 \pm 0.34

- 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 3.1. 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.

5.2 Computational Efficiency (Q2)

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. 4, and the cluster configuration in Table 4².

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

²GPU specs: Tesla P100-PCI-E-16G

Table 4: Distributed cluster configuration.

Cluster configuration	Parameter Server	Worker
# Instances	4	100
# CPU	28 cores	440 cores
# GPU	-	25 cards
Memory (GB)	40	1000

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.

6 CONCLUSION

In this paper, we proposed Multi-IPW and Multi-DR CVR estimators for industrial recommender system. Both CVR estimators aim to counter the inherent issues in practice: 1) selection bias, and 2) data sparsity. 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, the idea can be generalized to debiasing CTR estimation by exploiting the sequential pattern “item pool \rightarrow exposure \rightarrow click”.

REFERENCES

- [1] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W Bruce Croft. 2018. Unbiased learning to rank with unbiased propensity estimation. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. ACM, 385–394.
- [2] Peter C Austin. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research* 46, 3 (2011), 399–424.

- [3] Elias Bareinboim, Jin Tian, and Judea Pearl. 2014. Recovering from selection bias in causal and statistical inference. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*.
- [4] Arnaud De Mittyenaere, Bénédicte Le Grand, Boris Golden, and Fabrice Rossi. 2014. Reducing offline evaluation bias in recommendation systems. In *23rd annual Belgian-Dutch Conference on Machine Learning (Benelearn 2014)*. Bruxelles, Belgium, 55–62. <https://hal.archives-ouvertes.fr/hal-01017734>
- [5] Miroslav Dudík, John Langford, and Lihong Li. 2011. Doubly robust policy evaluation and learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning (ICML '11)*. Omnipress, Madison, WI, USA, 1097–1104.
- [6] Craig K Enders. 2010. *Applied missing data analysis*. Guilford press.
- [7] Mehrdad Farajtabar, Yinlam Chow, and Mohammad Ghavamzadeh. 2018. More robust doubly robust off-policy evaluation. In *Proceedings of the 35th International Conference on Machine Learning (Proceedings of Machine Learning Research)*, Jennifer Dy and Andreas Krause (Eds.), Vol. 80. PMLR, Stockholm, Sweden, 1447–1456. <http://proceedings.mlr.press/v80/farajtabar18a.html>
- [8] Tom Fawcett. 2006. An introduction to ROC analysis. *Pattern Recognition Letters* 27, 8 (2006), 861–874.
- [9] Ronald A Fisher. 1956. *Statistical methods and scientific inference*. Hafner Publishing Co.
- [10] Guy Hadash, Oren Sar Shalom, and Rita Osadchy. 2018. Rank and rate: multi-task learning for recommender systems. In *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, 451–454.
- [11] Keisuke Hirano, Guido W Imbens, and Geert Ridder. 2003. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71, 4 (2003), 1161–1189.
- [12] Jiayuan Huang, Arthur Gretton, Karsten Borgwardt, Bernhard Schölkopf, and Alex J Smola. 2007. Correcting sample selection bias by unlabeled data. In *Advances in Neural Information Processing Systems*. 601–608.
- [13] Guido W Imbens and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- [14] Kuang-chih Lee, Burkay Orten, Ali Dasdan, and Wentong Li. 2012. Estimating conversion rate in display advertising from past performance data. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 768–776.
- [15] RJA Little and DB Rubin. 2002. *Statistical analysis with missing data*. Wiley. New York (2002).
- [16] Roderick JA Little and Donald B Rubin. 2019. *Statistical analysis with missing data*. Vol. 793. John Wiley & Sons.
- [17] Xiao Ma, Liqin Zhao, Guan Huang, Zhi Wang, Zelin Hu, Xiaoqiang Zhu, and Kun Gai. 2018. 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)*. Association for Computing Machinery, New York, NY, USA, 1137–1140. <https://doi.org/10.1145/3209978.3210104>
- [18] Karthika Mohan, Judea Pearl, and Jin Tian. 2013. Graphical models for inference with missing data. In *Advances in Neural Information Processing Systems*. 1277–1285.
- [19] Yabo Ni, Dan Ou, Shichen Liu, Xiang Li, Wenwu Ou, Anxiang Zeng, and Luo Si. 2018. 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*. ACM, 596–605.
- [20] Donald B Rubin. 1976. Inference and missing data. *Biometrika* 63, 3 (1976), 581–592.
- [21] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as treatments: Debiasing learning and evaluation. In *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48 (ICML '16)*. JMLR.org, 1670–1679.
- [22] Harald Steck. 2010. 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*. ACM, 713–722.
- [23] Aaron Van den Oord, Sander Dieleman, and Benjamin Schrauwen. 2013. Deep content-based music recommendation. In *Advances in Neural Information Processing Systems*. 2643–2651.
- [24] Karel Vermeulen and Stijn Vansteelandt. 2015. Bias-reduced doubly robust estimation. *J. Amer. Statist. Assoc.* 110, 511 (2015), 1024–1036.
- [25] Hongwei Wang, Fuzheng Zhang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. 2019. Multi-task feature learning for knowledge graph enhanced recommendation. In *The World Wide Web Conference*. ACM, 2000–2010.
- [26] Jizhe Wang, Pipei Huang, Huan Zhao, Zhibo Zhang, Binqiang Zhao, and Dik Lun Lee. 2018. Billion-scale commodity embedding for e-commerce recommendation in alibaba. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 839–848.
- [27] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2019. Doubly robust joint learning for recommendation on data missing not at random. *Proceedings of Machine Learning Research* (2019).
- [28] Gary M Weiss. 2004. Mining with rarity: a unifying framework. *ACM Sigkdd Explorations Newsletter* 6, 1 (2004), 7–19.
- [29] Anne Whitehead and John Whitehead. 1991. A general parametric approach to the meta-analysis of randomized clinical trials. *Statistics in Medicine* 10, 11 (1991), 1665–1677.
- [30] Yuriko Yamaguchi, Mimpei Morishita, Youichi Inagaki, Reyn Nakamoto, Jianwei Zhang, Junichi Aoi, and Shinsuke Nakajima. 2016. 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*. 42–49.
- [31] Wenhao Zhang, Ramin Ramezani, and Arash Naeim. 2019. WOTBoost: Weighted oversampling technique in boosting for imbalanced learning. *arXiv e-prints*, Article arXiv:1910.07892 (Oct 2019), arXiv:1910.07892 pages. arXiv:cs.LG/1910.07892
- [32] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM.
- [33] Han Zhu, Junqi Jin, Chang Tan, Fei Pan, Yifan Zeng, Han Li, and Kun Gai. 2017. Optimized cost per click in taobao display advertising. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2191–2200.