Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random

Xiaojie Wang [1] **Rui Zhang** [1] **Yu Sun** [2] **Jianzhong Qi** [1]

Abstract

In recommender systems, usually the ratings of a user to most items are missing. A critical problem in recommendation learning is that, in reality, the missing ratings are often missing not at random (MNAR). MNAR ratings make it difficult to accurately predict the ratings and unbiasedly estimate the performance of rating prediction. Recent approaches use imputed errors to recover the prediction errors for missing ratings, or weight observed ratings with the propensities of being observed. These approaches can still be severely biased in performance estimation or suffer from the variance of the propensities. We propose a principled approach to overcome these limitations. First, we propose an estimator that integrates the imputed errors and propensities in a doubly robust way to obtain unbiased performance estimation and alleviate the effect of the propensity variance. Based on this estimator, we propose joint learning of rating prediction and error imputation to achieve good performance guarantees. Extensive experiments show that our approach outperforms the state-of-the-art on four real-world datasets.

1 Introduction

Users' preferences to items in recommender systems are often represented as binary or multi-scaled ratings (Zhang et al., 2017). Ratings are often sparse, i.e., only the ratings to a small portion of items are observed, whereas the ratings to most of the items are missing (Wang et al., 2018b). Most studies aim at recommending the items that users may like based on such sparse ratings (Bell et al., 2007).

Existing studies often assume that the missing ratings are missing at random (MAR), but usually this assumption does not hold and the missing ratings are *missing not at random* (MNAR). For example, a recent study in song recommendation shows that the probability of a rating being missing depends on the rating's value (Marlin et al., 2007). Recent studies also show that correctly adopting the MNAR assumption helps improve the quality of recommended items (Wang et al., 2018a). Hence, we aim to address the recommendation problem based on the MNAR assumption. Ratings are MNAR largely because users are free to choose what items to rate. Users normally rate an item that they like, and thus the ratings of a lower value are more likely to be missing. In other words, the *propensities*, i.e., the probabilities of different ratings being observed, are not the same.

MNAR ratings make it difficult to learn a prediction model that aims at accurately predicting the (true) rating of a user to an item (Lim et al., 2015). The common practice of using only observed ratings yields suboptimal prediction models because the observed ratings are not a representative sample of all ratings – whether observed or missing. MNAR ratings also make it difficult to correctly estimate the performance of a prediction model (Steck, 2011). The performance is typically defined as the prediction inaccuracy: the average of prediction errors (e.g. squared difference between a predicted rating and the true rating) for all ratings (Salakhutdinov et al., 2007). Given MNAR ratings, averaging over only observed ratings can be severely *biased*: it can over- or under-estimate the prediction inaccuracy by a large amount, a.k.a, the *bias* (Little & Rubin, 2014).

Two recent approaches address the MNAR problem: (1) The error-imputation-based (EIB) approach computes an *imputed error*, i.e., an estimated value of the prediction error, for each missing rating (Steck, 2013). (2) The inverse-propensity-scoring (IPS) approach inversely weights the prediction error for each observed rating with the propensity of observing that rating (Schnabel et al., 2016). The EIB approach often has a large bias due to imputation inaccuracy, i.e., deviations of the imputed errors from the prediction errors (Dudík et al., 2011). Such inaccuracy will also be propagated into training a prediction model and increase the prediction inaccuracy. The IPS approach often suffers from the high variance of the propensities (Thomas & Brunskill, 2016). The propensities, once inversed, can cause training

[1]University of Melbourne, Parkville, Victoria, Australia. [2]Twitter, San Francisco, California, United States of America. Correspondence to: Rui Zhang <rui.zhang@unimelb.edu.au>.

Proceedings of the 36^{th} *International Conference on Machine Learning*, Long Beach, California, PMLR 97, 2019.

losses to oscillate and lead to a poor generalization ability.

We propose a principled approach to overcome these limitations. First, we propose a doubly robust estimator of the prediction inaccuracy. The estimator corrects the deviations of the imputed errors, inversely weighted with the propensities, for observed ratings. We prove that the estimator has a desired property called *double robustness*: the capability to remain unbiased if either the imputed errors or propensities are accurate. Because of this property, our estimator reduces the bias in estimating the prediction inaccuracy and alleviates the effect of the large propensity variance. A challenge of using this estimator for learning a prediction model is to prevent the imputation inaccuracy from increasing the prediction inaccuracy. We analyze this issue by deriving a generalization bound and then propose jointly learning of a prediction model and an imputation model to address this challenge. The imputation model learns to accurately estimate the prediction errors made by the prediction model, while the prediction model learns from the imputation model to reduce the prediction errors in itself. In this way, the prediction and imputation models mutually regularize each other to reduce both prediction and imputation inaccuracies.

The contributions of this paper are summarized as follows.

- We propose a doubly robust estimator for unbiased performance estimation in recommendation and theoretically analyze the bias and tail bound of the estimator.
- Based on the doubly robust estimator, we propose a novel approach to jointly learn rating prediction and error imputation, which enjoys a rigorous performance guarantee.
- We conduct extensive experiments and the results show that the proposed approach achieves up to a 12% drop in the prediction inaccuracy compared with the state-of-the-art.

2 Preliminaries

Let $\mathcal{U} = \{u_1, ..., u_N\}$ be a set of users, $\mathcal{I} = \{i_1, ..., i_M\}$ a set of items, and $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ the set of all user-item pairs. Let $\mathbf{R} \in \mathbb{R}^{N \times M}$ be a *true rating matrix* where each entry $r_{u,i}$ is the true rating of user u to item i. A recommendation method learns a prediction model that aims to predict the true ratings and recommends items with the highest predicted ratings (Ricci et al., 2010). Let $\hat{\mathbf{R}} \in \mathbb{R}^{N \times M}$ be a *prediction matrix* where each entry $\hat{r}_{u,i}$ is a predicted rating computed by the prediction model. If we have a fully observed true rating matrix $\mathbf{R}^f$, the *prediction inaccuracy* $\mathcal{P}$ of the prediction model can be measured by metrics such as mean absolute error (MAE) or mean square error (MSE)

$$\mathcal{P} = \mathcal{P}(\hat{\mathbf{R}}, \mathbf{R}^f) = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} e_{u,i}\,, \tag{1}$$

where $e_{u,i} = |\hat{r}_{u,i} - r_{u,i}|$ or $e_{u,i} = (\hat{r}_{u,i} - r_{u,i})^2$ is the *prediction error* for MAE or MSE, respectively. Let $\mathbf{O} \in \{0,1\}^{N \times M}$ be an *indicator matrix* where each entry $o_{u,i}$ is an observation indicator: $o_{u,i} = 1$ if the true rating $r_{u,i}$ is observed, and $o_{u,i} = 0$ if the true rating $r_{u,i}$ is missing. Let $\mathbf{R}^o$ and $\mathbf{R}^m$ be the set of the observed and missing entries in the true rating matrix. Given the observed ratings $\mathbf{R}^o$, the rating prediction problem aims to learn the prediction model that minimizes the prediction inaccuracy $\mathcal{P}$.

Since most entries in the true rating matrix are often missing, a naive (N) estimator estimates the prediction inaccuracy by averaging the prediction errors for observed ratings

$$\mathcal{E}_{\mathrm{N}} = \mathcal{E}_{\mathrm{N}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{O}|} \sum_{u,i \in \mathcal{O}} e_{u,i}\,,$$

where $\mathcal{O} = \{(u,i) | u,i \in \mathcal{D}, o_{u,i} = 1\}$ is the set of user-item pairs for the observed ratings. If the missing ratings $\mathbf{R}^m$ are missing at random (MAR), the naive estimator is *unbiased*: the expectation of its estimation over all the possible instances of $\mathbf{O}$ is exactly the same as the prediction inaccuracy, i.e., $\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{N}}] = \mathcal{P}$. MAR means that the probability of observing an instance of the indicator matrix only depends on the observed ratings $p(\mathbf{O}|\mathbf{R}, \mathbf{X}) = p(\mathbf{O}|\mathbf{R}^o)$, where $\mathbf{X}$ are all the other factors that affect the indicator matrix besides the true rating matrix $\mathbf{R}$ (Liang et al., 2016). In some cases, MAR does not hold and the missing ratings are missing not at random (MNAR), e.g., the probability of a rating being missing depends on its value (Marlin et al., 2007). In such cases, the naive estimator can have a large *bias*: a large difference between the prediction inaccuracy and the expectation of its estimation over $\mathbf{O}$, i.e., $|\mathcal{P} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{N}}]| \gg 0$.

To reduce the bias, an error-imputation-based (EIB) estimator uses an imputation model to compute imputed errors, i.e., estimated values of the prediction errors (Steck, 2010). An imputation model, called *heuristic imputation*, computes the *imputed error* $\hat{e}_{u,i} = \omega|\hat{r}_{u,i} - \gamma|$ or $\hat{e}_{u,i} = \omega(\hat{r}_{u,i} - \gamma)^2$ for MAE or MSE, where ω and γ are hyper-parameters (Steck, 2010). Using the imputed errors for missing ratings together with the prediction errors for observed ratings, the EIB estimator estimates the prediction inaccuracy with

$$\mathcal{E}_{\mathrm{EIB}} = \mathcal{E}_{\mathrm{EIB}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} (o_{u,i} e_{u,i} + (1 - o_{u,i}) \hat{e}_{u,i})\,,$$

where we call $\delta_{u,i} = e_{u,i} - \hat{e}_{u,i}$ the *error deviation*. Alternatively, Schnabel et al. (2016) learn the *propensity* $p_{u,i} = P(o_{u,i} = 1)$, i.e., the probability of observing the true rating $r_{u,i}$ by, e.g., naive bayes. They use the learned propensity $\hat{p}_{u,i}$ to inversely weight each prediction error for observed ratings and define an inverse-propensity-scoring (IPS) estimator that estimates the prediction inaccuracy with

$$\mathcal{E}_{\mathrm{IPS}} = \mathcal{E}_{\mathrm{IPS}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}}\,.$$

When the imputed errors are accurate, i.e., $\delta_{u,i} = 0$ for $u,i \in \mathcal{D}$, the EIB estimator is unbiased no matter whether the missing ratings are MAR or MNAR (Vermeulen & Vansteelandt, 2015). Similarly, the IPS estimator is also unbiased when the propensities are accurate, i.e., $\Delta_{u,i} = \frac{\hat{p}_{u,i} - p_{u,i}}{\hat{p}_{u,i}} = 0$ for $u, i \in \mathcal{D}$. However, the EIB estimator usually has a large bias in practice due to the *imputation inaccuracy* which can be measured by metrics such as MSE

$$\mathcal{I} = \mathcal{I}(\hat{\mathbf{R}}, \mathbf{R}^f) = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} \delta_{u,i}^2 \, .$$

On the other hand, the IPS estimator often suffers from a high variance (Gilotte et al., 2018). Such variance can be reduced by a self-normalized inverse propensity scoring (SNIPS) estimator (Swaminathan & Joachims, 2015b)

$$\mathcal{E}_{\text{SNIPS}} = \mathcal{E}_{\text{SNIPS}}(\hat{\mathbf{R}}, \mathbf{R}^o) = \left(\sum_{u,i \in \mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}} \right)^{-1} \sum_{u,i \in \mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}} \, .$$

We propose to use both imputed errors and propensities to overcome the limitations of the EIB and IPS approaches.

3 Doubly Robust Estimator

A straightforward idea of using both imputed errors and propensities is to combine the EIB and IPS estimators as $\mathcal{E}_{\text{SC}} = \lambda \mathcal{E}_{\text{EIB}} + (1-\lambda)\mathcal{E}_{\text{IPS}}$. A weakness of such a linear combination is that even when the propensities are accurate, the bias increases if the imputed errors become less accurate. We observe that this weakness can be addressed by designing an estimator in a doubly robust way such that the bias remains zero even with deteriorated imputed errors. The key idea is to correct the error deviation $\delta_{u,i}$ for observed ratings and inversely weight the corrections with the propensity $\hat{p}_{u,i}$ to consider the MNAR effect. Following this idea, we propose such an estimator, called a *doubly robust* (DR) estimator, to estimate the prediction inaccuracy.

Given the imputed errors $\hat{\mathbf{E}} = \{\hat{e}_{u,i} | u, i \in \mathcal{D}\}$ and learned propensities $\hat{\mathbf{P}} = \{\hat{p}_{u,i} | u, i \in \mathcal{D}\}$, the DR estimator estimates the prediction inaccuracy $\mathcal{P}$ with

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

The DR estimator augments the IPS estimator with the following low-variance term (Seaman et al., 2018)

$$\mathcal{E}_{\text{DR}} - \mathcal{E}_{\text{IPS}} = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} \frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}} \, .$$

Note that the expectation of this term over $\mathbf{O}$ is equal to zero given accurate propensities (see appendix for proofs).

Figure 1 illustrates the advantage of the DR estimator. At the top of Figure 1, we have the true rating matrix $\mathbf{R}$ and

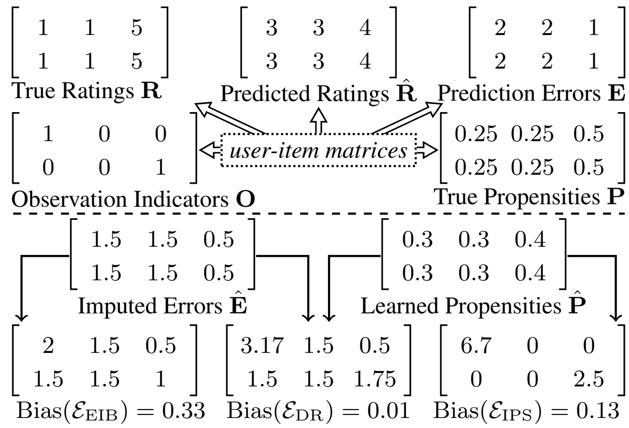


Figure 1: Our DR estimator ($\mathcal{E}_{\text{DR}}$) of the prediction inaccuracy has a smaller bias than the EIB ($\mathcal{E}_{\text{EIB}}$) and IPS ($\mathcal{E}_{\text{IPS}}$) estimators.

prediction matrix $\hat{\mathbf{R}}$. If the true rating matrix is fully observed, we have the prediction error matrix $\mathbf{E} = |\mathbf{R} - \hat{\mathbf{R}}|$ from MAE at the top right. We compute the prediction inaccuracy $\mathcal{P} = 1.67$ by averaging over all entries of the prediction error matrix. For illustration purpose, we make a MNAR assumption: only one of the four ratings of value 1 (the top-left entry) and one of the two ratings of value 5 (the bottom-right entry) are observed (indicated by the indicator matrix $\mathbf{O}$ at the middle left). Hence, the true propensities of observing the ratings of value 1 and 5 are 0.25 and 0.5 (shown in the true propensities $\mathbf{P}$ at the middle right). We use the heuristic imputation with $\omega = 1$ and $\gamma = 4.5$, which produces the imputed errors of 1.5 and 0.5 for the ratings of value 1 and 5 (shown in the imputed errors $\hat{\mathbf{E}}$). Hence, the bias of the EIB estimator in this example is $\text{Bias}(\mathcal{E}_{\text{EIB}}) = |\mathcal{P} - \mathcal{E}_{\text{EIB}}| = 0.33$. We perturb the true propensities by 20%, which gives us the learned propensities of 0.3 and 0.4 for the ratings of value 1 and 5 (shown in the learned propensities $\hat{\mathbf{P}}$). The bias of the IPS estimator is $\text{Bias}(\mathcal{E}_{\text{IPS}}) = |\mathcal{P} - \mathcal{E}_{\text{IPS}}| = 0.13$. By using both imputed errors and propensities, the bias of the simple combination is $\text{Bias}(\mathcal{E}_{\text{SC}}) \in [0.13, 0.33]$, whereas the bias of the DR estimator is the smallest $\text{Bias}(\mathcal{E}_{\text{DR}}) = |\mathcal{P} - \mathcal{E}_{\text{DR}}| = 0.01$.

We formally derive the bias of the DR estimator as follows.

Lemma 3.1 (Bias of DR Estimator)**.** *Given imputed errors* $\hat{\mathbf{E}}$ *and learned propensities* $\hat{\mathbf{P}}$ *with* $\hat{p}_{u,i} > 0$ *for all user-item pairs, the bias of the DR estimator is*

$$\text{Bias}(\mathcal{E}_{\text{DR}}) = \frac{1}{|\mathcal{D}|} \left| \sum_{u,i \in \mathcal{D}} \Delta_{u,i} \delta_{u,i} \right| . \tag{2}$$

Proof. By definition, the bias of the DR estimator is

$$\text{Bias}(\mathcal{E}_{\text{DR}}) = |\mathcal{P} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}]| \, .$$

The second term on the right hand side can be expanded as

$$\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}] = \frac{1}{|\mathcal{D}|} \sum_{u,i \in \mathcal{D}} \left(\hat{e}_{u,i} + \frac{p_{u,i}\delta_{u,i}}{\hat{p}_{u,i}} \right) . \tag{3}$$

Table 1: Bias of the EIB, IPS, and DR estimators.

$\mathcal{E}_{\text{EIB}}$	$\mathcal{E}_{\text{IPS}}$	$\mathcal{E}_{\text{DR}}$
$\left\lvert \sum_{u,i\in\mathcal{D}} \frac{(1-p_{u,i})\delta_{u,i}}{\lvert\mathcal{D}\rvert} \right\rvert$	$\left\lvert \sum_{u,i\in\mathcal{D}} \frac{\Delta_{u,i}e_{u,i}}{\lvert\mathcal{D}\rvert} \right\rvert$	$\left\lvert \sum_{u,i\in\mathcal{D}} \frac{\Delta_{u,i}\delta_{u,i}}{\lvert\mathcal{D}\rvert} \right\rvert$

We treat the prediction and imputed errors as constants when taking the expectation since $\mathbf{O}$ does not result from any prediction or imputation models (Schnabel et al., 2016). Subtracting Eq. 3 from Eq. 1 yields the stated results. □

We summarize the bias of the EIB, IPS, and DR estimators in Table 1. If either $\delta_{u,i} \approx 0$ or $\Delta_{u,i} \approx 0$, the DR estimator is close to the prediction inaccuracy, whereas the EIB estimator requires $\delta_{u,i} \approx 0$ and the IPS estimator requires $\Delta_{u,i} \approx 0$. Moreover, if $\delta_{u,i} \approx 0$ and $\Delta_{u,i} \ll 1 - p_{u,i}$, the DR estimator is less biased than the EIB estimator. Similarly, if $\Delta_{u,i} \approx 0$ and $\delta_{u,i} \ll e_{u,i}$, the DR estimator is less biased than the IPS estimator. Thus, the DR estimator effectively takes advantage of both imputed errors and propensities for less biased estimation of the prediction inaccuracy.

We formally describe *double robustness* as follows. The proof substitutes either $\delta_{u,i} = 0$ or $\Delta_{u,i} = 0$ into the bias of the DR estimator in Eq. 2 (see appendix for details).

Corollary 3.1 (Double Robustness)**.** *The DR estimator is unbiased when either imputed errors $\hat{\mathbf{E}}$ or learned propensities $\hat{\mathbf{P}}$ are accurate for all user-item pairs.*

We next analyze the tail bound of the DR estimator (see appendix for proofs). Following existing work (Schnabel et al., 2016), we assume that the indicator matrix $\mathbf{O}$ contains independent random variables and each one $o_{u,i}$ follows a Bernoulli distribution with probability $p_{u,i}$.

Lemma 3.2 (Tail Bound of DR Estimator)**.** *Given imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$, for any prediction matrix $\hat{\mathbf{R}}$, with probability $1-\eta$, the deviation of the DR estimator from its expectation has the following tail bound*

$$\left\lvert \mathcal{E}_{\text{DR}} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}] \right\rvert \leq \sqrt{\frac{\log\left(\frac{2}{\eta}\right)}{2\lvert\mathcal{D}\rvert^2} \sum_{u,i\in\mathcal{D}} \left(\frac{\delta_{u,i}}{\hat{p}_{u,i}}\right)^2}.$$

We further present the following corollary to compare the tail bound of the DR and IPS estimators (see appendix for proofs). The corollary indicates that the DR estimator will almost surely have a lower tail bound than the IPS estimator when the imputed errors do not deviate from the prediction errors by a large amount, e.g., $\lvert\delta_{u,i}\rvert \leq e_{u,i}$. We will also empirically show such results in the experiments.

Corollary 3.2 (Tail Bound Comparison)**.** *Suppose imputed errors $\hat{\mathbf{E}}$ are such that $0 \leq \hat{e}_{u,i} \leq 2e_{u,i}$ for $u,i \in \mathcal{D}$, then for any learned propensities $\hat{\mathbf{P}}$, the tail bound of the DR estimator will be lower than that of the IPS estimator.*

4 Joint Learning Approach

A challenge of using the DR estimator for recommendation learning is to prevent the error deviations caused by the imputation model from worsening the training of the prediction model. We derive a generalization bound to analyze this issue and observe how the error deviations affect the prediction inaccuracy of a prediction model trained with the DR estimator. Based on the observation, we propose a joint learning approach to address this challenge.

Given observed ratings $\mathbf{R}^o$, we obtain the optimal prediction matrix $\hat{\mathbf{R}}^{\ddagger}$ by minimizing the estimated prediction inaccuracy by the DR estimator over a hypothesis space $\mathcal{H}$ of prediction matrices (Sugiyama & Kawanabe, 2012)

$$\hat{\mathbf{R}}^{\ddagger} = \underset{\hat{\mathbf{R}}\in\mathcal{H}}{\operatorname{argmin}} \left\{ \mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}, \mathbf{R}^o) \right\}.$$

We prove that the prediction inaccuracy of the optimal prediction matrix has the following generalization bound (see appendix for proofs). For ease of presentation, we use the same superscript to denote the correspondence between an error deviation and a prediction matrix (e.g., $\delta^{\ddagger}_{u,i}$ and $\hat{\mathbf{R}}^{\ddagger}$).

Theorem 4.1 (Generalization Bound)**.** *For any finite hypothesis space $\mathcal{H}$ of prediction matrices, with probability $1-\eta$, the prediction inaccuracy $\mathcal{P}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^f)$ of the optimal prediction matrix using the DR estimator with imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ has the upper bound*

$$\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^o) + \underbrace{\sum_{u,i\in\mathcal{D}} \frac{\lvert\Delta_{u,i}\delta^{\ddagger}_{u,i}\rvert}{\lvert\mathcal{D}\rvert}}_{\text{Bias Term}} + \underbrace{\sqrt{\frac{\log\left(\frac{2\lvert\mathcal{H}\rvert}{\eta}\right)}{2\lvert\mathcal{D}\rvert^2} \sum_{u,i\in\mathcal{D}} \left(\frac{\delta^{\S}_{u,i}}{\hat{p}_{u,i}}\right)^2}}_{\text{Variance Term}},$$

where $\delta^{\S}_{u,i}$ is the error deviation corresponding to the prediction matrix $\hat{\mathbf{R}}^{\S} = \operatorname{argmax}_{\hat{\mathbf{R}}^h\in\mathcal{H}} \left\{ \sum_{u,i\in\mathcal{D}} \left(\frac{\delta^h_{u,i}}{\hat{p}_{u,i}}\right)^2 \right\}$.

We can see that the generalization bound contains a bias term and a variance term, both of which are positively correlated with the magnitude of the error deviation $\lvert\delta_{u,i}\rvert$. Hence, the generalization bound indicates that we can guarantee a low prediction inaccuracy if the error deviations are of a small magnitude, i.e., the imputation inaccuracy is also low. To minimize both prediction and imputation inaccuracies, we propose joint learning of a prediction model $f_\theta(\boldsymbol{x}_{u,i})$ and an imputation model $g_\phi(\boldsymbol{x}_{u,i})$. The prediction model $\hat{r}_{u,i} = f_\theta(\boldsymbol{x}_{u,i})$, parameterized by θ, aims to accurately predict the true rating $r_{u,i}$ given a vector $\boldsymbol{x}_{u,i}$ encoding all the features of user u and item i. We train the prediction model by minimizing the estimated prediction inaccuracy by the DR estimator and use the training loss

$$\mathcal{L}_{\text{r}}(\theta, \phi) = \sum_{u,i\in\mathcal{D}} \left(\hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}} \right) + \upsilon \lVert\theta\rVert^2_F,$$

Algorithm 1 Alternating Training for Joint Learning

input: observed ratings $\mathbf{R}^o$ and learned propensities $\hat{\mathbf{P}}$
while *stopping criteria is not satisfied* **do**
  for *number of steps for training the imputation model* **do**
    Sample a batch of user-item pairs $\{(u_j, i_j)\}_{j=1}^{J}$ from $\mathcal{O}$
    Update ϕ by descending along the gradient $\nabla_\phi \mathcal{L}_\mathrm{e}(\theta, \phi)$
  end for
  for *number of steps for training the prediction model* **do**
    Sample a batch of user-item pairs $\{(u_k, i_k)\}_{k=1}^{K}$ from $\mathcal{D}$ [2]
    Update θ by descending along the gradient $\nabla_\theta \mathcal{L}_\mathrm{r}(\theta, \phi)$
  end for
end while

where $e_{u,i} = (f_\theta(\boldsymbol{x}_{u,i}) - r_{u,i})^2$, $\hat{e}_{u,i} = (f_\theta(\boldsymbol{x}_{u,i}) - g_\phi(\boldsymbol{x}_{u,i}) - \perp(f_\theta(\boldsymbol{x}_{u,i})))^2$ with $\perp$ the operator that sets the gradient of the operand to zero so $\nabla_\theta \perp(f_\theta(\boldsymbol{x}_{u,i})) = 0$ and $\perp(f_\theta(\boldsymbol{x}_{u,i})) = f_\theta(\boldsymbol{x}_{u,i})$ [1], $\upsilon \geq 0$, and $\|\cdot\|_F^2$ is the Frobenius norm. Meanwhile, we also learn the parameters of the imputation model, which we call *imputation learning*. The imputation model $\hat{e}_{u,i} = g_\phi(\boldsymbol{x}_{u,i})$, parameterized by ϕ, aims to accurately estimate the prediction error $e_{u,i}$ given the feature vector $\boldsymbol{x}_{u,i}$. We train the imputation model by minimizing the squared deviations of the imputed errors from the prediction errors and use the training loss

$$\mathcal{L}_\mathrm{e}(\theta, \phi) = \sum_{u,i \in \mathcal{O}} \frac{(\hat{e}_{u,i} - e_{u,i})^2}{\hat{p}_{u,i}} + \nu \|\phi\|_F^2 ,$$

where $e_{u,i} = r_{u,i} - f_\theta(\boldsymbol{x}_{u,i})$, $\hat{e}_{u,i} = g_\phi(\boldsymbol{x}_{u,i})$, and $\nu \geq 0$. We inversely weight the squared deviation for each observed rating with the propensity $\hat{p}_{u,i}$ to consider the MNAR effect. We compute the prediction error $e_{u,i}$ by the difference (instead of the absolute or squared difference), so the imputation model can learn to distinguish whether a predicted rating is larger or smaller than the true rating. We alternate between training the prediction and imputation models via minbatch stochastic gradient descent (Wang et al., 2018d). We summarize the alternating training process in Alg. 1. Note that Theorem 4.1 assumes fixed imputed errors and provides performance guarantees on the inner loop of Alg. 1.

5 Experiments

First, we compare our joint learning approach with existing rating prediction approaches to show its effectiveness and flexibility. Next, we create a synthetic dataset to study the bias and standard deviation of our doubly robust estimator.

5.1 Inaccuracy in Rating Prediction Problem

Dataset. Unbiased estimation of the prediction inaccuracy needs MAR ratings (Marlin & Zemel, 2009). To our knowledge, two real-world datasets have MAR ratings as follows.

[1]This operator $\perp$ has been implemented in scientific libraries, e.g., TensorFlow as `stop_gradient` and PyTorch as `detach`.

[2]Due to the sparsity of observed ratings, we decrease the probability of unobserved ratings being sampled in the experiments.

Table 2: Inaccuracy of rating prediction on MAR test ratings.

	COAT		YAHOO	
	MAE	MSE	MAE	MSE
MF	0.920	1.257	1.154	1.891
PMF	0.903	1.239	1.103	1.709
AutoRec	0.900	1.238	0.984	1.438
Gaussian-VAE	0.893	1.220	0.963	1.381
CPT-v	0.969	1.441	0.770	1.115
MF-HI	0.922	1.261	1.158	1.905
MF-MNAR	0.884	1.214	1.177	2.175
MF-IPS	0.860	1.093	0.810	0.989
MF-JL	0.866	1.136	0.899	1.256
MF-DR-JL	**0.778**	**0.990**	**0.747**	**0.966**

* MF-JL and MF-DR-JL are the proposed approaches.

1. COAT has 4,640 MAR and 6,960 MNAR ratings of 290 users to 300 coats (Schnabel et al., 2016).
2. YAHOO has 54,000 MAR and 311,704 MNAR ratings of 15,400 users to 1,000 songs (Marlin & Zemel, 2009).

We also conduct experiments on the following two real-world datasets that have only MNAR ratings.

1. AMAZON has 1,000,086 MNAR ratings of 33,326 users to 21,901 television shows (He & McAuley, 2016).
2. MOVIE has 10,000,054 MNAR ratings of 71,567 users to 10,681 movies (Harper & Konstan, 2016).

Experiment Setup. Following prior work (Schnabel et al., 2016), we use MNAR ratings for training and MAR ratings for testing on COAT and YAHOO. This allows us to unbiasedly evaluate the capability of an approach to debias learning on biased data. Since AMAZON and MOVIE do not have MAR ratings, we randomly split the MNAR ratings into a training set (90%) and a test set (10%) (Zheng et al., 2016). We use 5-fold cross-validation to set the hyper-parameters of all approaches. We use the methods proposed by Schnabel et al. (2016) to learn propensities as follows. On COAT, we use logistic regression based on all pairs of user features (e.g., gender) and item features (e.g., color). We cross-validate logistic regression to maximize the log-likelihood of observing the MNAR ratings. YAHOO, AMAZON, and MOVIE do not have user and item features, which makes logistic regression not applicable. Hence, we set aside 5% of the test ratings and use naive bayes to learn propensities.

Prediction Inaccuracy. We call approaches that explicitly deal with MNAR ratings and those that do not *debiasing* and *biased* approaches, respectively. We compare with the following debiasing approaches: **CPT-v** (Marlin & Zemel, 2009), **MF-HI** (Steck, 2010), **MF-MNAR** (Hernández-Lobato et al., 2014), and **MF-IPS** (Schnabel et al., 2016). We also compare with the following biased approaches: **MF** (Koren et al., 2009), **PMF** (Mnih & Salakhutdinov, 2008), **AutoRec** (Sedhain et al., 2015), and **Gaussian-VAE** (Liang et al., 2018). We call our joint learning ap-

Table 3: Inaccuracy of rating prediction on MNAR test ratings.

	AMAZON		MOVIE	
	MSE	MSE-SNIPS	MSE	MSE-SNIPS
MF	0.949	0.931	0.803	0.793
PMF	0.969	0.911	0.824	0.773
AutoRec	0.900	0.887	0.782	0.776
Gaussian-VAE	0.874	0.861	0.770	0.765
CPT-v	1.277	1.236	1.235	1.180
MF-HI	0.964	0.935	0.812	0.803
MF-MNAR	0.943	0.913	0.803	0.764
MF-IPS	0.956	0.924	0.819	0.780
MF-JL	**0.868**	0.851	**0.767**	0.756
MF-DR-JL	0.871	**0.844**	0.782	**0.745**

* MF-JL and MF-DR-JL are the proposed approaches.

proach using MF (Koren et al., 2009) to implement the prediction and imputation models **MF-DR-JL**. We call MF-DR-JL trained with uniform propensities $\hat{p}_{u,i} = 1$ **MF-JL**.

The prediction inaccuracy under MAE and MSE on COAT and YAHOO (datasets with MAR ratings) is shown in Table 2. Our MF-DR-JL performs the best under both metrics on both datasets, e.g., MF-DR-JL (0.778) outperforms MF-IPS (0.860) by 10% under MAE on COAT. Both imputed errors and propensities are critical to MF-DR-JL since removing either (MF-IPS or MF-JL) causes a significant increase in the prediction inaccuracy. In general, the debiasing approaches outperform the biased ones except that: (1) MF-MNAR performs worst on YAHOO. A possible explanation is that MF-MNAR makes several generative assumptions (e.g., hierarchical Gaussian priors) which may not hold on YAHOO. Unlike MF-MNAR, our MF-DR-JL does not make any generative assumptions and performs consistently well across datasets. (2) CPT-v performs worst on COAT. This is partly because CPT-v learns propensities by conditioning on the true ratings. The learned propensities are inaccurate because they can depend on user and item features. In contrast to CPT-v, MF-DR-JL is more robust to inaccurate propensities and learns more accurate propensities by regressing on all available user and item features. (3) MF-HI performs even worse than MF, which suggests that the heuristic imputation can be harmful for learning a prediction model.

The prediction inaccuracy under MSE and MSE-SNIPS on AMAZON and MOVIE is reported in Table 3. The results under MAE and MAE-SNIPS are similar to those in Table 3 and are put in the appendix. To simulate testing on MAR ratings, we compute MAE-SNIPS and MSE-SNIPS with the learned propensities (Swaminathan & Joachims, 2015a). We can see that our MF-JL and MF-DR-JL are competitive with or better than the other approaches. MF-JL outperforms MF-DR-JL under MSE, while opposite results are obtained under MSE-SNIPS. This is expected because uniform propensities are used by MF-JL and MSE, whereas learned propensities are used by MF-DR-JL and MSE-SNIPS. Overall, the debiasing approaches do not have clear advantages over the biased ones when tested on MNAR ratings.

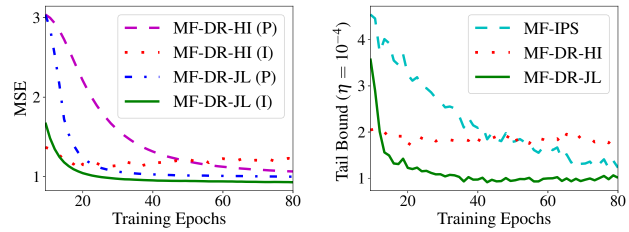
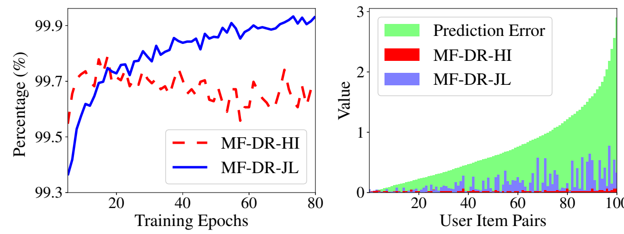


Figure 2: Advantage of imputation learning (best view in color).



(a) Percentage of user-item pairs that satisfy $0 \leq \hat{e}_{u,i} \leq 2e_{u,i}$.

(b) User-item pairs are sorted by the prediction error $e_{u,i}$.

Figure 3: Advantage of imputation learning (MF-DR-JL).

Error Imputation. We study the advantage of the imputation learning over the heuristic imputation.We call MF-DR-JL using the heuristic imputation **MF-DR-HI**. The left of Figure 2 shows the prediction ($\mathcal{P}$) and imputation ($\mathcal{I}$) inaccuracies under MSE against training epochs on COAT. We can see that MF-DR-JL has a lower prediction inaccuracy than MF-DR-HI because the imputation learning effectively reduces the imputation inaccuracy. We further use MAR ratings to estimate the tail bound ($\eta = 10^{-4}$) of estimators used at training. The right of Figure 2 shows the tail bound against training epochs on COAT. We can see that the tail bound at training MF-DR-JL is lower than that at training MF-IPS, while that at training MF-DR-HI remains high.

Recall that when $0 \leq \hat{e}_{u,i} \leq 2e_{u,i}$, the tail bound of the DR estimator is lower than that of the IPS estimator. We use MAR ratings to compute the percentage of user-item pairs that satisfy the condition $0 \leq \hat{e}_{u,i} \leq 2e_{u,i}$ under MSE. Figure 3a shows the percentage against training epochs on COAT. We can see that most user-item pairs ($> 99\%$) satisfy the condition. The percentage increases until reaching around 99.9% when the imputation learning is used. Figure 3b shows the values of the prediction and imputed errors for 100 user-item pairs, randomly selected and sorted by the prediction errors, at epoch 80 on COAT. We can see that the imputed errors that violate the condition have a very small magnitude. The tail bounds are dominated by the imputed errors that satisfy the condition and have a large value. Besides, the heuristic imputation results in quite small imputed errors that do not well correlate with the prediction errors.

Propensity Variance. We study how the propensity variance affects the learning process. Let $\hat{p}_l = \max_{\mathcal{D}}\{\hat{p}_{u,i}\}$

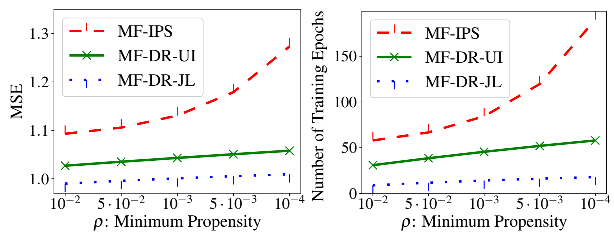


Figure 4: Learning with varying levels of propensity variance.

and $\hat{p}_s = \min_{\mathcal{D}}\{\hat{p}_{u,i}\}$, we use propensities rescaled by $\tilde{p}_{u,i} = \hat{p}_l - \frac{(\hat{p}_l-\rho)(\hat{p}_l-\hat{p}_{u,i})}{\hat{p}_l-\hat{p}_s}$ for learning. The variance of the rescaled propensities increases as $\rho = \min_{\mathcal{D}}\{\tilde{p}_{u,i}\}$ goes to 0. We call MF-DR-JL using uniform propensities to train the imputation model **MF-DR-UI**. Figure 4 shows MSE (left) and number of training epochs took to reach an MSE of 1.5 (right) against varying levels of propensity variance on COAT. We can see that MF-DR-JL has a lower MSE and takes less training epochs than MF-IPS at all levels of propensity variance. MF-DR-JL also has a slower increase in MSE and number of training epochs than MF-IPS as the propensity variance grows. These results suggest that the proposed approach is less affected by the high variance of the propensities. By comparing MF-DR-JL with MF-DR-UI, we can also see the benefit of using the propensities to train the imputation model on MNAR ratings.

Model Implementation. We study how different model implementations affect the prediction inaccuracy. We compare with two other biased approaches: **FM** (Rendle, 2010) and **NFM** (He & Chua, 2017). FM generalizes MF to work with any user and item features and NFM extends FM to model high-order user-item interactions. We also compare with two other debiasing approaches: **FM-IPS** and **NFM-IPS** (Schnabel et al., 2016). We call our joint learning approach using FM (or NFM) to implement the prediction and imputation models **FM-DR-JL** (or **NFM-DR-JL**). We call FM-DR-JL (or NFM-DR-JL) trained with $\hat{p}_{u,i} = 1$ **FM-JL** (or **NFM-JL**). The prediction inaccuracy under MAE and MSE on COAT and YAHOO is reported in Table 4. We can see that our approaches benefit from improved model implementations. For example, NFM-DR-JL (0.756) outperforms MF-DR-JL (0.778) by 3% under MAE on COAT.

5.2 Bias of Prediction Inaccuracy Estimation

Dataset. Computing the prediction inaccuracy requires a fully observed true rating matrix, which does not exist in any real-world datasets. Hence, we create such a true rating matrix on YAHOO as follows (Schnabel et al., 2016). First, we use MF (Koren et al., 2009) to complete the partial true rating matrix but MF gives unrealistically high ratings to all items. We therefore adjust the completed true rating matrix to match a more realistic rating distribution $[v_1, v_2, v_3, v_4, v_5]$, estimated from the MAR ratings (Marlin & Zemel, 2009), for values of 1 to 5. We achieve this by sorting the matrix entries in ascending order, assigning a value of 1 to the lowest v_1 fraction of the matrix entries, assigning a value of 2 to the next v_2 fraction, and so on.

Table 4: Inaccuracy of rating prediction on MAR test ratings.

	COAT		YAHOO	
	MAE	MSE	MAE	MSE
FM	0.911	1.252	1.154	1.891
NFM	0.888	1.218	1.001	1.488
FM-IPS	0.853	1.086	0.810	0.989
NFM-IPS	0.832	1.065	0.798	0.979
FM-JL	0.859	1.129	1.032	1.528
NFM-JL	0.838	1.114	1.016	1.509
FM-DR-JL	0.775	0.986	0.747	0.966
NFM-DR-JL	**0.756**	**0.967**	**0.736**	**0.957**

* The bottom four rows show the proposed approaches.

Experiment Setup. For fair comparison, we use the following five prediction matrices used by Schnabel et al. (2016).

1. ONE: We randomly select 137,800 entries (i.e., the same number of the true ratings with a value of 5) with a value of 1 and set their values to 5 in the true rating matrix.
2. FOUR: We randomly select 137,800 entries with a value of 4 and set their values to 5 in the true rating matrix.
3. ROT: The predicted rating $\hat{r}_{u,i} = r_{u,i} - 1$ if the true rating $r_{u,i} \geq 2$. Otherwise, the predicted rating $\hat{r}_{u,i} = 5$.
4. SKEW: The predicted rating $\hat{r}_{u,i}$ is sampled from $\mathcal{N}(\mu = r_{u,i}, \sigma = \frac{6-r_{u,i}}{2})$ and is clipped to $[0, 6]$.
5. CRS: The predicted rating $\hat{r}_{u,i} = 4$ if the true rating $r_{u,i} \geq 4$. Otherwise, the predicted rating $\hat{r}_{u,i} = 3$.

We model the MNAR effect where higher ratings are more likely to be observed on YAHOO (Schnabel et al., 2016). We define the true propensity $p_{u,i} = p$ when the true rating $r_{u,i} = 5$, and $p_{u,i} = p\alpha^{\min(4,6-r_{u,i})}$ otherwise. We obtain an instance of $\mathbf{O}$ by sampling for each entry an observation indicator $o_{u,i}$ from a Bernoulli distribution with probability $p_{u,i}$. We control the MNAR effect by varying α and p and set their values so that the expected rating sparsity is 5%. To produce a rating distribution of observed ratings that reasonably matches the observed MNAR rating distribution on YAHOO, we set $\alpha = 0.5$ in all experiments.

We simulate error imputation of varying inaccuracies using the heuristic imputation. By varying $\omega \in [0, 1]$, we vary the inaccuracy of imputed errors: as ω goes to 0, the imputed errors usually become less accurate. We compute the values of ω and γ by $\sum_{\mathcal{D}} \hat{e}_{u,i} = \sum_{\mathcal{D}} \frac{o_{u,i}e_{u,i}}{\hat{p}_{u,i}}$ and $\gamma = \frac{1}{|\mathcal{D}|}\sum_{\mathcal{D}} \frac{o_{u,i}r_{u,i}}{\hat{p}_{u,i}}$, unless otherwise stated.

We also simulate propensity estimation of varying inaccuracies. We obtain the propensities by $\frac{1}{\hat{p}_{u,i}} = \frac{1-\beta}{p_{u,i}} + \frac{\beta}{p_e}$ and $p_e = \frac{1}{|\mathcal{D}|}\sum_{\mathcal{D}} o_{u,i}$. We vary the inaccuracy of the propensities by varying $\beta \in [0, 1]$. We obtain similar results when varying β and set $\beta = 0.5$, unless otherwise specified.

Table 5: Bias and standard deviation in terms of percentage over the prediction inaccuracy under MSE. DR is the proposed estimator.

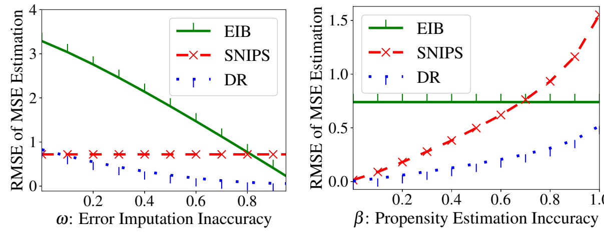
	EIB	IPS	SNIPS	NCIS	DR
ONE	22.8±1.8	20.7±1.8	20.7±1.8	26.0±1.7	**9.9±0.9**
FOUR	64.5±1.7	66.8±1.8	66.8±1.8	84.0±1.8	**24.1±0.6**
ROT	18.4±0.3	18.5±0.3	18.5±0.2	23.1±0.2	**10.3**±0.2
SKEW	15.7±0.5	14.8±0.7	14.9±0.5	17.8±0.4	**10.1±0.3**
CRS	18.6±0.3	16.1±0.5	16.2±0.3	20.7±0.2	**9.0±0.1**

Bias and Standard Deviation. We compare our **DR** estimator with the **EIB**, **IPS**, **SNIPS**, and **NCIS** (the propensities are clipped to $(0, \frac{1}{40}])$ estimators (Gilotte et al., 2018). We first use the estimators to estimate the prediction inaccuracy of the five prediction matrices for the given level of error imputation and propensity estimation inaccuracies. We compute the bias and standard deviation of the estimators using 50 instances of the indicator matrix. The results in terms of the percentage over the prediction inaccuracy under MSE are shown in Table 5 (see appendix for the results under MAE). We can see that the DR estimator achieves the smallest bias with the lowest standard deviation.

In the last set of experiments, we study how the inaccuracies of imputed errors and propensities affect the bias of the estimators. We compute the bias with root mean squared error (RMSE) over the five prediction matrices and 50 instances of the indicator matrix (Schnabel et al., 2016).

Inaccurate Imputation. We study the bias of the EIB and DR estimators when varying the inaccuracy of imputed errors $\omega \in [0, 1]$. The DR estimator only uses a small number (50) of MAR ratings to learn propensities via naive bayes. As shown in Figure 5a, the DR estimator is consistently less biased than the EIB estimator, especially when the imputed errors are severely inaccurate, e.g., $\omega = 0$.

Inaccurate Propensities. We study the bias of the EIB and SNIPS (the best baseline) estimators when varying the inaccuracy of propensities $\beta \in [0, 1]$. Figure 5b shows that the bias of the SNIPS estimator increases rapidly as β goes to 1. The bias of the DR estimator increases gradually because the DR estimator uses relatively accurate imputed errors to reduce the bias caused by the deteriorated propensities.



(a) Varying ω. (b) Varying β.

Figure 5: Bias of varying imputation and propensity inaccuracies.

6 Related Work

Doubly robust approaches are first proposed in statistical inference (Benkeser et al., 2017; Morgan & Winship, 2015) and are used in online advertising (Chan et al., 2010; Raeder et al., 2012) and reinforcement learning (Dudík et al., 2011; Farajtabar et al., 2018). These studies differ from ours in that: (1) They do not employ doubly robust estimators as training losses while we do, which is challenging since using an imputation model may hurt the prediction inaccuracy; (2) They only have rewards that reveal part of target variables while we have target variables to learn, which gives our approach a rigorous performance guarantee. Doubly robust approaches are also applied to recommender systems for bias-variance tradeoff in offline evaluation (Gilotte et al., 2018). This recent work bases its theoretical analyses on the assumption that the propensities are accurate. We do not require such a strong assumption and theoretically analyze how inaccurate propensities impact the bias and tail bound.

Most debiasing approaches for recommendation rely on a single missing data model (Marlin et al., 2007; Steck, 2013). We integrate two missing data models in a doubly robust way such that our approach is less affected by the mis-specification of the missing data models.

Prior studies show the benefit of joint learning of classification and imputation models (Van Esbroeck et al., 2014). Unlike these studies, we weight the joint learning with the propensities to make our approach robust to inaccuracy imputation models. Compared with the expectation maximization (EM) approach (Dempster et al., 1977), our approach also alternates between imputing missing data and updating model parameters but does not require to take expectation.

Also related to the studied rating prediction problem is item ranking problem, which aims to optimize ranking metrics (He et al., 2017; Wang et al., 2018c). These two problems are of great importance and widely studied in industry and academia (He et al., 2016; Zhao et al., 2019). The accuracy of item ranking can also suffer from various biases (Ai et al., 2018). These studies are orthogonal to our work and our approach can be adapted to item ranking by, e.g., building on pointwise learning-to-rank (Liu et al., 2009).

7 Conclusions

We proposed a principled approach to handle missing not at random data for recommendation. First, we proposed a doubly robust estimator, which achieves double robustness by using both imputed errors and propensities to estimate the prediction inaccuracy. Next, we proposed a joint learning approach, which learns rating prediction and error imputation jointly to guarantee a low prediction inaccuracy at inference time. We conducted extensive experiments on four real-world datasets. The results showed that our approach outperforms the state-of-the-art in rating prediction. The results also showed that the proposed estimator significantly reduces the bias of estimating the prediction inaccuracy.

Acknowledgement

This work is supported by Australian Research Council Future Fellowship Project FT120100832 and Discovery Project DP180102050. We would like to thank the anonymous reviewers for their insightful comments on the paper.

References

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

Bell, R., Koren, Y., and Volinsky, C. Chasing $1,000,000: How we won the netflix progress prize. *Statistical Computing and Graphics*, 2007.

Benkeser, D., Carone, M., Laan, M. V. D., and Gilbert, P. Doubly robust nonparametric inference on the average treatment effect. *Biometrika*, 2017.

Chan, D., Ge, R., Gershony, O., Hesterberg, T., and Lambert, D. Evaluating online ad campaigns in a pipeline: causal models at scale. In *SIGKDD*, 2010.

Dempster, A. P., Laird, N. M., and Rubin, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B*, 1977.

Dudík, M., Langford, J., and Li, L. Doubly robust policy evaluation and learning. In *ICML*, 2011.

Farajtabar, M., Chow, Y., and Ghavamzadeh, M. More robust doubly robust off-policy evaluation. In *ICML*, 2018.

Gilotte, A., Calauzènes, C., Nedelec, T., Abraham, A., and Dollé, S. Offline a/b testing for recommender systems. In *WSDM*, 2018.

Harper, F. M. and Konstan, J. A. The movielens datasets: History and context. *TiiS*, 2016.

He, R. and McAuley, J. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *WWW*, 2016.

He, X. and Chua, T.-S. Neural factorization machines for sparse predictive analytics. In *SIGIR*, 2017.

He, X., Zhang, H., Kan, M.-Y., and Chua, T.-S. Fast matrix factorization for online recommendation with implicit feedback. In *SIGIR*, 2016.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T.-S. Neural collaborative filtering. In *WWW*, 2017.

Hernández-Lobato, J. M., Houlsby, N., and Ghahramani, Z. Probabilistic matrix factorization with non-random missing data. In *ICML*, 2014.

Koren, Y., Bell, R., and Volinsky, C. Matrix factorization techniques for recommender systems. *Computer*, 2009.

Liang, D., Charlin, L., McInerney, J., and Blei, D. M. Modeling user exposure in recommendation. In *WWW*, 2016.

Liang, D., Krishnan, R. G., Hoffman, M. D., and Jebara, T. Variational autoencoders for collaborative filtering. In *WWW*, 2018.

Lim, D., McAuley, J., and Lanckriet, G. Top-n recommendation with missing implicit feedback. In *RecSys*, 2015.

Little, R. J. and Rubin, D. B. *Statistical analysis with missing data*. John Wiley & Sons, 2014.

Liu, T.-Y. et al. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 2009.

Marlin, B., Zemel, R. S., Roweis, S., and Slaney, M. Collaborative filtering and the missing at random assumption. In *UAI*, 2007.

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

Mnih, A. and Salakhutdinov, R. R. Probabilistic matrix factorization. In *NeurIPS*, 2008.

Morgan, S. L. and Winship, C. *Counterfactuals and causal inference*. Cambridge University Press, 2015.

Raeder, T., Stitelman, O., Dalessandro, B., Perlich, C., and Provost, F. Design principles of massive, robust prediction systems. In *SIGKDD*, 2012.

Rendle, S. Factorization machines. In *ICDM*, 2010.

Ricci, F., Rokach, L., Shapira, B., and Kantor, P. B. *Recommender Systems Handbook*. Springer-Verlag, 2010.

Salakhutdinov, R., Mnih, A., and Hinton, G. Restricted boltzmann machines for collaborative filtering. In *ICML*, 2007.

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

Seaman, S. R., Vansteelandt, S., et al. Introduction to double robust methods for incomplete data. *Statistical Science*, 2018.

Sedhain, S., Menon, A. K., Sanner, S., and Xie, L. Autorec: Autoencoders meet collaborative filtering. In *WWW*, 2015.

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

Steck, H. Item popularity and recommendation accuracy. In *RecSys*, 2011.

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

Sugiyama, M. and Kawanabe, M. *Machine learning in non-stationary environments: Introduction to covariate shift adaptation*. MIT press, 2012.

Swaminathan, A. and Joachims, T. Counterfactual risk minimization: Learning from logged bandit feedback. In *ICML*, 2015a.

Swaminathan, A. and Joachims, T. The self-normalized estimator for counterfactual learning. In *NeurIPS*, 2015b.

Thomas, P. and Brunskill, E. Data-efficient off-policy policy evaluation for reinforcement learning. In *ICML*, 2016.

Van Esbroeck, A., Singh, S. P., Rubinfeld, I., and Syed, Z. Evaluating trauma patients: Addressing missing covariates with joint optimization. In *AAAI*, 2014.

Vermeulen, K. and Vansteelandt, S. Bias-reduced doubly robust estimation. *Journal of the American Statistical Association*, 2015.

Wang, M., Gong, M., Zheng, X., and Zhang, K. Modeling dynamic missingness of implicit feedback for recommendation. In *NeurIPS*, 2018a.

Wang, X., Qi, J., Ramamohanarao, K., Sun, Y., Li, B., and Zhang, R. A joint optimization approach for personalized recommendation diversification. In *PAKDD*, 2018b.

Wang, X., Wen, J.-R., Dou, Z., Sakai, T., and Zhang, R. Search result diversity evaluation based on intent hierarchies. *TKDE*, 2018c.

Wang, X., Zhang, R., Sun, Y., and Qi, J. Kdgan: Knowledge distillation with generative adversarial networks. In *NeurIPS*, 2018d.

Zhang, S., Yao, L., and Sun, A. Deep learning based recommender system: A survey and new perspectives. *arXiv preprint arXiv:1707.07435*, 2017.

Zhao, Y., Qi, J., and Zhang, R. Cbhe: Corner-based building height estimation for complex street scene images. In *WWW*, 2019.

Zheng, Y., Tang, B., Ding, W., and Zhou, H. A neural autoregressive approach to collaborative filtering. In *ICML*, 2016.

Supplementary Material of "Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random"

Xiaojie Wang [1] **Rui Zhang** [1] **Yu Sun** [2] **Jianzhong Qi** [1]

1 Proofs of Lemmas and Theorems

Lemma 1.1 (Zero-Expectation Term)**.** *Given any imputed errors $\hat{\mathbf{E}}$, suppose learned propensities $\hat{\mathbf{P}}$ are accurate, the expectation of the term $\mathcal{E}_{\mathrm{DR}} - \mathcal{E}_{\mathrm{IPS}}$ over all the possible instances of the indicator matrix $\mathbf{O}$ is equal to zero.*

Proof. When the learned propensities are accurate, we have $\Delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. Therefore, we can compute the expectation of the term $\mathcal{E}_{\mathrm{DR}} - \mathcal{E}_{\mathrm{IPS}}$ over all the possible instances of the indicator matrix $\mathbf{O}$ by

$$\begin{aligned}
&\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}} - \mathcal{E}_{\mathrm{IPS}}] \\
&= \mathbb{E}_{\mathbf{O}}\left[\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&= \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{\mathbf{O}}\left[\frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&= \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{o_{u,i}}\left[\frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&= \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \frac{(\hat{p}_{u,i} - p_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}, \\
&= \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\hat{e}_{u,i}, \\
&= 0.
\end{aligned}$$

This completes the proof. □

Lemma 1.2 (Bias of EIB Estimator)**.** *Given imputed errors $\hat{\mathbf{E}}$, the bias of the EIB estimator is given by*

$$\mathrm{Bias}\left(\mathcal{E}_{\mathrm{EIB}}\right) = \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}}(1 - p_{u,i})\delta_{u,i}\right|.$$

Proof. According to the definition of the bias, we can derive the bias of the EIB estimator as follows

$$\begin{aligned}
&\mathrm{Bias}\left(\mathcal{E}_{\mathrm{EIB}}\right) \\
&= \left|\mathcal{P} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{EIB}}]\right|, \\
&= \left|\mathcal{P} - \mathbb{E}_{\mathbf{O}}\left[\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} o_{u,i}e_{u,i} + (1 - o_{u,i})\hat{e}_{u,i}\right]\right|, \\
&= \left|\mathcal{P} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{o_{u,i}}[o_{u,i}e_{u,i} + (1 - o_{u,i})\hat{e}_{u,i}]\right|, \\
&= \left|\mathcal{P} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} (p_{u,i}e_{u,i} + (1 - p_{u,i})\hat{e}_{u,i})\right|, \\
&= \left|\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} e_{u,i} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} (p_{u,i}e_{u,i} + (1 - p_{u,i})\hat{e}_{u,i})\right|, \\
&= \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} (1 - p_{u,i})\delta_{u,i}\right|,
\end{aligned}$$

which completes the proof. □

Corollary 3.1 (Double Robustness)**.** *The DR estimator is unbiased when either imputed errors $\hat{\mathbf{E}}$ or learned propensities $\hat{\mathbf{P}}$ are accurate for all user-item pairs.*

Proof. On one hand, when the imputed errors are accurate, we have $\delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. In such case, we can compute the bias of the DR estimator by

$$\begin{aligned}
&\mathrm{Bias}(\mathcal{E}_{\mathrm{DR}}) \\
&= \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\delta_{u,i}\right|, \\
&= \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i} \cdot 0\right|, \\
&= 0\,.
\end{aligned}$$

On the other hand, when the learned propensities are accurate, we have $\Delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. In this case, we can

[1]University of Melbourne, Parkville, Victoria, Australia. [2]Twitter, San Francisco, California, United States of America. Correspondence to: Rui Zhang <rui.zhang@unimelb.edu.au>.

Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

compute the bias of the DR estimator by

$$\begin{aligned}&\text{Bias}(\mathcal{E}_{\text{DR}})\\&=\frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}}\Delta_{u,i}\delta_{u,i}\right|,\\&=\frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}}0\cdot\delta_{u,i}\right|,\\&=0\,.\end{aligned}$$

In both cases, the bias of the DR estimator is zero, which means that the expectation of the DR estimator over all the possible instances of $\mathbf{O}$ is exactly the same as the prediction inaccuracy. This completes the proof. ☐

Lemma 3.2 (Tail Bound of DR Estimator)**.** *Given imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$, for any prediction matrix $\hat{\mathbf{R}}$, with probability $1-\eta$, the deviation of the DR estimator from its expectation has the following tail bound*

$$\left|\mathcal{E}_{\text{DR}}-\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}]\right|\leq\sqrt{\frac{\log\left(\frac{2}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}}{\hat{p}_{u,i}}\right)^2}.$$

Proof. To keep the notation uncluttered, we define random variable $\ell_{u,i}$ that equals

$$\ell_{u,i}=\hat{e}_{u,i}+\frac{o_{u,i}(e_{u,i}-\hat{e}_{u,i})}{\hat{p}_{u,i}}\,.$$

Since we assume that each observation indicator $o_{u,i}$ follows a Bernoulli distribution with probability $p_{u,i}$, we can rewrite the random variable $\ell_{u,i}$ as follows

$$\begin{cases}P(\ell_{u,i}=\hat{e}_{u,i}+\rho_{u,i})=p_{u,i}\,,\\P(\ell_{u,i}=\hat{e}_{u,i})=1-p_{u,i}\,,\end{cases}$$

where $\rho_{u,i}$ is given by

$$\rho_{u,i}=\frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}}=\frac{\delta_{u,i}}{\hat{p}_{u,i}}\,.$$

We can see that the random variable $\ell_{u,i}$ takes value in the interval $[\hat{e}_{u,i},\hat{e}_{u,i}+\rho_{u,i}]$ of size $\rho_{u,i}$ with probability 1. Recall that we assume that the observation indicators $\{o_{u,i}|u,i\in\mathcal{D}\}$ are independent random variables, so the random variables $\{\ell_{u,i}|u,i\in\mathcal{D}\}$ are also independent. Therefore, according to the Hoeffding's inequality, for any $\tilde{\epsilon}>0$, we have the following inequality

$$P\left(\left|\sum_{u,i}\ell_{u,i}-\mathbb{E}_{\mathbf{O}}\left[\sum_{u,i}\ell_{u,i}\right]\right|\geq\tilde{\epsilon}\right)\leq 2\exp\left(\frac{-2\tilde{\epsilon}^2}{\sum_{u,i}\rho_{u,i}^2}\right).$$

Here, the three summations are summed over all user-item pairs $u,i\in\mathcal{D}$. Setting $\tilde{\epsilon}=\epsilon|\mathcal{D}|$ ($\tilde{\epsilon}>0\Leftrightarrow\epsilon>0$) in the above inequality and simplifying the inequality with the definition of the DR estimator yield

$$P\left(\left|\mathcal{E}_{\text{DR}}-\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}]\right|\geq\epsilon\right)\leq 2\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i}\rho_{u,i}^2}\right).$$

Setting the right hand side of the inequality to η and solving for ϵ complete the proof. ☐

Corollary 3.2 (Tail Bound Comparison)**.** *Suppose imputed errors $\hat{\mathbf{E}}$ are such that $0\leq\hat{e}_{u,i}\leq 2e_{u,i}$ for $u,i\in\mathcal{D}$, then for any learned propensities $\hat{\mathbf{P}}$, the tail bound of the DR estimator will be lower than that of the IPS estimator.*

Proof. To simplify the notation, we define a constant C as

$$C=\frac{\log\left(\frac{2}{\eta}\right)}{2|\mathcal{D}|^2}\,.$$

Then, we can derive the following inequalities

$$\begin{aligned}&0\leq\hat{e}_{u,i}\leq 2e_{u,i}\text{ for }u,i\in\mathcal{D}\,,\\&\Rightarrow\delta_{u,i}^2\leq e_{u,i}^2\text{ for }u,i\in\mathcal{D}\,,\\&\Rightarrow\sqrt{C\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}}{\hat{p}_{u,i}}\right)^2}\leq\sqrt{C\sum_{u,i\in\mathcal{D}}\left(\frac{e_{u,i}}{\hat{p}_{u,i}}\right)^2}\,.\end{aligned}$$

In the last inequality, the left hand side is the tail bound of the DR estimator and the right hand side is the tail bound of the IPS estimator (Schnabel et al., 2016). This completes the proof. ☐

Theorem 4.1 (Generalization Bound)**.** *For any finite hypothesis space $\mathcal{H}$* [1] *of prediction matrices, with probability $1-\eta$, the prediction inaccuracy $\mathcal{P}(\hat{\mathbf{R}}^{\ddagger},\mathbf{R}^f)$ of the optimal prediction matrix using the DR estimator with imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ has the upper bound*

$$\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}^{\ddagger},\mathbf{R}^o)+\underbrace{\sum_{u,i\in\mathcal{D}}\frac{|\Delta_{u,i}\delta_{u,i}^{\ddagger}|}{|\mathcal{D}|}}_{\textit{Bias Term}}+\underbrace{\sqrt{\frac{\log\left(\frac{2|\mathcal{H}|}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}^{\S}}{\hat{p}_{u,i}}\right)^2}}_{\textit{Variance Term}},$$

where $\delta_{u,i}^{\S}$ is the error deviation corresponding to the prediction matrix $\hat{\mathbf{R}}^{\S}=\text{argmax}_{\hat{\mathbf{R}}^h\in\mathcal{H}}\left\{\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}^h}{\hat{p}_{u,i}}\right)^2\right\}$.

Proof. To simplify notation, let $\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}})=\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}},\mathbf{R}^o)$. Note that we can rewrite the finite hypothesis space with $\mathcal{H}=\{\hat{\mathbf{R}}^1,...,\hat{\mathbf{R}}^{|\mathcal{H}|}\}$. We define $\rho_{u,i}^h$ as follows

$$\rho_{u,i}^h=\frac{e_{u,i}^h-\hat{e}_{u,i}^h}{\hat{p}_{u,i}}=\frac{\delta_{u,i}^h}{\hat{p}_{u,i}}\,,$$

[1] For infinite hypothesis spaces, a similar generalization bound can be derived using covering numbers (Anthony & Bartlett, 2009) or other related measures (Maurer & Pontil, 2009).

where $e^h_{u,i}$ and $\hat{e}^h_{u,i}$ are the prediction and imputed errors corresponding to a prediction matrix $\hat{\mathbf{R}}^h \in \mathcal{H}$, respectively. Similarly, we define $\rho^\S_{u,i}$ as follows

$$\rho^\S_{u,i} = \frac{e^\S_{u,i} - \hat{e}^\S_{u,i}}{\hat{p}_{u,i}} = \frac{\delta^\S_{u,i}}{\hat{p}_{u,i}},$$

where $e^\S_{u,i}$ and $\hat{e}^\S_{u,i}$ are the prediction and imputed errors corresponding to the prediction matrix $\hat{\mathbf{R}}^\S \in \mathcal{H}$, respectively. By making the arguments of uniform convergence and union bound, for any $\epsilon > 0$, we have

$$
\begin{aligned}
& P\Big(\Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^\ddagger) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^\ddagger)]\Big| \le \epsilon\Big) \ge 1 - \eta, \\
& \Leftarrow P\left(\max_{\hat{\mathbf{R}}^h \in \mathcal{H}} \Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \le \epsilon\right) \ge 1 - \eta \\
& \quad \text{(by the argument of uniform convergence)}, \\
& \Leftrightarrow P\left(\bigcup_{\hat{\mathbf{R}}^h \in \mathcal{H}} \Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \ge \epsilon\right) < \eta, \\
& \Leftarrow \sum_{h=1}^{|\mathcal{H}|} P\Big(\Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \ge \epsilon\Big) < \eta \\
& \quad \text{(by the argument of union bound)}, \\
& \Leftarrow \sum_{h=1}^{|\mathcal{H}|} 2\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i\in\mathcal{D}}\{\rho^h_{u,i}\}^2}\right) < \eta \\
& \quad \text{(by Hoeffding's inequality in Lemma 3.2)}, \\
& \Leftarrow 2|\mathcal{H}|\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i\in\mathcal{D}}\{\rho^\S_{u,i}\}^2}\right) < \eta \\
& \quad \text{(by the definition of the prediction matrix } \hat{\mathbf{R}}^\S).
\end{aligned}
$$

We solve the inequality in the last line for ϵ and obtain, with probability $1 - \eta$, the following inequality

$$
\begin{aligned}
& \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^\ddagger)] - \mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^\ddagger) \\
& \le \sqrt{\frac{\log\left(\frac{2|\mathcal{H}|}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta^\S_{u,i}}{\hat{p}_{u,i}}\right)^2}.
\end{aligned} \tag{1}
$$

Given the optimal prediction matrix $\hat{\mathbf{R}}^\ddagger$, imputed errors $\hat{\mathbf{E}}^\ddagger = \{\hat{e}^\ddagger_{u,i} | u,i \in \mathcal{D}\}$, and learned propensities $\hat{\mathbf{P}}$, the bias of the DR estimator can be upper bounded as follows

$$
\begin{aligned}
& \mathcal{P}(\hat{\mathbf{R}}^\ddagger, \mathbf{R}^f) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^\ddagger)] \\
& \le \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\delta^\ddagger_{u,i}\right|, \\
& \le \sum_{u,i\in\mathcal{D}} \frac{|\Delta_{u,i}\delta^\ddagger_{u,i}|}{|\mathcal{D}|}.
\end{aligned} \tag{2}
$$

After adding the two inequalities in Eq. 1 and Eq. 2, we can rearrange the terms to obtain the stated results. □

2 Additional Experiment Results

Table 1 reports the results of rating prediction measured by MAE and MAE-SNIPS on AMAZON and MOVIE.

Table 1: Inaccuracy of rating prediction on MNAR test ratings.

	AMAZON		MOVIE	
	MAE	MAE-SNIPS	MAE	MAE-SNIPS
MF	0.764	0.761	0.745	0.743
PMF	0.767	0.764	0.754	0.748
AutoRec	0.759	0.755	0.743	0.737
Gaussian-VAE	0.730	0.727	0.733	0.728
CPT-v	1.001	0.991	0.956	0.939
MF-HI	0.773	0.770	0.749	0.742
MF-MNAR	0.747	0.739	0.741	0.727
MF-IPS	0.768	0.759	0.752	0.738
MF-JL	**0.721**	0.720	**0.694**	0.693
MF-DR-JL	0.725	**0.717**	0.703	**0.689**

* MF-JL and MF-DR-JL are the proposed approaches.

Table 2 shows the bias and standard deviation of the EIB, IPS, SNIPS, NCIS, and DR estimators in terms of the percentage over the prediction inaccuracy under MAE on the synthetic dataset (see Section 5.2 for details).

Table 2: Bias and standard deviation in terms of percentage over the prediction inaccuracy under MAE. DR is the proposed estimator.

	EIB	IPS	SNIPS	NCIS	DR
ONE	21.7±1.7	20.5±1.8	20.5±1.8	25.8±1.6	**9.8±0.8**
FOUR	66.9±1.7	66.8±1.8	66.8±1.8	84.0±1.8	**24.1±0.6**
ROT	12.7±0.2	12.7±0.4	12.8±0.2	16.0±0.2	**7.1±0.1**
SKEW	10.9±0.3	10.3±0.7	10.5±0.3	12.2±0.2	**7.1**±0.2
CRS	13.5±0.2	12.4±0.6	12.5±0.4	16.5±0.2	**6.9±0.1**

References

Anthony, M. and Bartlett, P. L. *Neural network learning: Theoretical foundations*. cambridge university press, 2009.

Maurer, A. and Pontil, M. Empirical bernstein bounds and sample variance penalization. In *COLT*, 2009.

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