Discriminative-Invariant Representation Learning for Unbiased Recommendation

anonymouse author

Abstract

Selection bias hinders recommendation models from learning unbiased user preference. Recent works empirically reveal that pursuing invariant user and item representation across biased and unbiased data is crucial for counteracting selection bias. However, our theoretical analysis reveals that simply optimizing representation invariance is insufficient for addressing the selection bias — recommendation performance is bounded by both representation invariance and discriminability. Worse still, current invariant representation learning methods in recommendation neglect even hurt the representation discriminability due to data sparsity and label shift. In this light, we propose a new Discriminative-Invariant Representation Learning framework for unbiased recommendation, which incorporates label-conditional clustering and prior-guided contrasting into conventional invariant representation learning to mitigate the impact of data sparsity and label shift, respectively. We conduct extensive experiments on three real-world datasets, validating the rationality and effectiveness of the proposed framework.

1 Introduction

Recommender systems (RS) make personalized recommendation by predicting the user preference over items, which achieves great success in a variety of applications [Chen *et al.*, 2020]. Historical feedback (*e.g.,* like or dislike) is indispensable for learning user preference, which is typically collected from previous recommendation strategies. Consequently, the historical feedback is subject to selection bias [Saito and Nomura, 2022], *i.e.,* the probability of being observed is unevenly distributed over user-item pairs. Blindly fitting such biased data will result in biased user preference [Chen *et al.*, 2020] and notorious issues such as unfairness and filter bubble [Huang *et al.*, 2022]. It is thus crucial to counteract the selection bias in user preference learning.

The key lies in aligning the distribution of user-item pairs in the biased data to the unbiased data with feedback collected under random exposure. Existing methods achieve the goal by minimizing distribution discrepancy, which are in two main categories at data level [Swaminathan and Joachims, 2015; Schnabel *et al.*, 2016; Wang *et al.*, 2019; Guo *et al.*, 2021] and representation level [Saito and Nomura, 2022; Wang *et al.*, 2022]. Bridging the discrepancy at data level is typically achieved by reweighting the user-item pairs according to the exposure probability (*a.k.a.* propensity score). These propensity-based methods typically suffer from unreliable propensity estimation [Saito and Nomura, 2022] and high variance [Guo *et al.*, 2021]. At the representation level, pursuing invariant user and item representations across different distributions is an emerging solution for minimizing distribution discrepancy. These invariant representation-based methods avoid the drawback of propensity scores, achieving promising empirical results [Wang *et al.*, 2022].

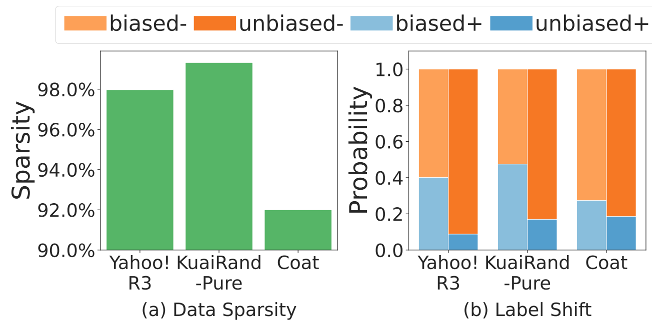

Figure 1: Data sparsity and label shift in Yahoo!R3, KuaiRand-Pure, and Coat. (a) Only a few feedback are observed in all user-item pairs. (b) The distribution of labels in biased data differs from that in unbiased data. Biased and unbiased denote biased data and unbiased data respectively while + and - denote positive feedback (*i.e.,* like) and negative feedback (*i.e.,* dislike).

Nevertheless, it remains unclear whether representation invariance is sufficient for unbiased recommendation. To bridge this research gap, we analyze the theoretical connection between the representation and recommendation performance based on the theory in unsupervised domain adaptation (UDA) [Ben-David *et al.*, 2010], finding that the recommendation performance is bounded by both the invariance of representation between the biased and unbiased data, and the discriminability of representation among user-item pairs with different labels. However, existing invariant representation-based methods typically assumes that the representation discriminability is sufficiently high and blindly optimize the invariance [Jiang *et al.*, 2022].

We argue that representation discriminability should be additionally optimized in recommendation due to the extreme data sparsity and severe label shift [Zhao *et al.*, 2019] between the biased and unbiased data as shown in Figure 1. The increase of data sparsity will inevitably limit the representation discriminability [Koren *et al.*, 2009], *i.e.,* the underlying assumption of existing invariant representation-based methods might be violated. Worse still, blindly optimizing the invariance will hurt the discriminability as the presence of label shift. This is because blindly optimizing the invariance will force the representation of some user-item pairs with different labels to be close to mitigate the discrepancy between label distributions. Therefore, optimizing the invariance of representation with consideration of both data sparsity and label shift has the potential of achieving better unbiased recommendation.

In this light, we propose a new *Discriminative-Invariant Representation Learning* (**DIRL**) framework for unbiased recommendation. Specifically, we equip the learning objective for pursuing invariant representation with two newly designed losses of label-conditional clustering and prior-guided contrasting to optimize both representation invariance and discriminability. Label-conditional clustering increases the distance among representation centroids of different labels to enhance the discriminability. Prior-guided contrasting restricts the distance between the predictions on the biased data and unbiased data according to the prior of label shift to avoid the elimination of representation distribution discrepancy from label shift. We implement DIRL based on a representative invariant representation learning method named adversarial distribution alignment [Ganin and Lempitsky, 2015; Tzeng *et al.*, 2015] and conduct evaluation on three real-world datasets. Extensive experimental results validate the rationality and effectiveness of the proposed framework.

The main contributions of this paper are as follows:

- We conduct theoretical analysis on the performance bound of unbiased recommendation, revealing the importance of both representation invariance and discriminability.
- We propose a new Discriminative-Invariant Representation Learning framework to address the selection bias issue in recommendation, which consists of two new losses to mitigate the impact of data sparsity and label shift.
- We conduct extensive experiments on three real-world datasets, validating the rationality and effectiveness of the proposed DIRL framework.

2 Analysis on Unbiased Recommendation

In this section, we first formulate the task of unbiased recommendation (Sec. 2.1). We then review recent invariant representation-based methods, uncovering the mystery of their effectiveness and identifying their limitations (Sec. 2.2). Finally, we empirically demonstrate the importance of boosting representation discriminability in RS (Sec. 2.3).

2.1 Task Formulation

We are given a RS with a user set $\mathcal{U}$ and an item set $\mathcal{I}$. Let u (or i) denote a user (or an item) in $\mathcal{U}$ (or $\mathcal{I}$). Let $f := \mathcal{U} \times \mathcal{I} \rightarrow \{0,1\}$ be the label function, indicating whether a user actually likes an item ($y = 1$) or not ($y = 0$). Historical feedback data can be formulated as a set of user-item pairs $\mathcal{D}_B := \{(u,i)_B^j\}_{j=1}^n$ drawn from a biased distribution $p_B(u,i)$ (*e.g.,* subject to previous recommendation policy) and their corresponding labels $f(\mathcal{D}_B) = \{r_B^j\}_{j=1}^n$. The task of unbiased recommendation can be formulated as follow: learning a recommendation model from the available biased data for capturing user preference and accordingly making a high-quality recommendation. Formally, the goal is to learn a function $h := \mathcal{U} \times \mathcal{I} \rightarrow \{0,1\}$ from biased data that approaches the ideal f over the test distribution:

$$\varepsilon_U(h,f) := E_{(u,i)\sim p_U}\left[l(h\left(u,i\right), f\left(u,i\right))\right], \quad (1)$$

where $p_U(u,i)$ denotes the unbiased distribution, which is often assumed as uniform (*i.e.,* random exposure); $l(.,.)$ denotes the selected error function between the prediction and the ground truth. Since labels for user-item pairs $\mathcal{D}_U := \{(u,i)_U^j\}_{j=1}^n$ drawn from $p_U(u,i)$ are not available, the model can be only trained on the biased data with optimizing the following empirical loss:

$$\hat{\varepsilon}_B(h,f) := E_{(u,i)\in D_B}\left[l(h\left(u,i\right), f\left(u,i\right))\right]. \quad (2)$$

As the distribution of the training dataset differs from the test, blindly fitting the data result in inferior performance and notorious issues like filter bubble[Huang *et al.*, 2022]. Thus, it is essential to develop a debiasing strategy for making great recommendations.

2.2 Analyses over Existing Invariant Representation-based Methods

To better understand the effectiveness and limitation of these methods, we first introduce the theorem in UDA [Ben-David *et al.*, 2010] for analyzing the generalized error bound of unbiased recommendation. In fact, we have:

Theorem 1. *(Generalized Error Bound.) Let $\mathcal{H}$ be a hypothesis space with VC-dimension d. With the probability of at least $1-\eta$, $\forall h \in \mathcal{H}$,*

$$\begin{aligned}\varepsilon_U(h,f) \leq \hat{\varepsilon}_B(h,f) + \frac{1}{2}d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_B,\mathcal{D}_U) \\ + \lambda^* + O(\sqrt{\frac{dlogn + log(1/\eta)}{n}})\end{aligned} \quad (3)$$

where $\mathcal{H}\Delta\mathcal{H}$-divergence $d_{\mathcal{H}\Delta\mathcal{H}}(\hat{\mathcal{D}}_B, \hat{\mathcal{D}}_U)$ measures the distribution discrepancy between biased and unbiased data. The ideal joint error $\lambda^* := min_{h\in\mathcal{H}}\varepsilon_B(h,f) + \varepsilon_U(h,f)$ is the optimal error that can be achieved by the hypotheses in $\mathcal{H}$ on both biased and unbiased data.

Following existing work [Ganin and Lempitsky, 2015; Zhang *et al.*, 2019], we can apply a representation function $g := \mathcal{U} \times \mathcal{I} \rightarrow \mathcal{Z}$ that maps the user-item space $\mathcal{U} \times \mathcal{I}$ to the representation space $\mathcal{Z}$ before passing through hypothesis h[1], which means we can analyze the bound from

[1]There is a little abuse of the notation h here.The domain of definition is changed.

the data level to the representation level. Accordingly, besides the empirical error and constant, the unbiased recommendation performance is also subjected to: 1) representation invariance, which pursues low representation distribution discrepancy between the biased and unbiased data ($d_{\mathcal{H}\Delta\mathcal{H}}(g(\mathcal{D}_B), g(\mathcal{D}_U))$). 2) representation discriminability, the ability for separating different labels by a supervised classifier trained over the representations in both biased and unbiased data [Chen *et al.*, 2019] (λ^*).

There is a type of method leveraging the extra regularizer to control the representation distribution discrepancy:

CVIB [Wang *et al.*, 2020]. The objective of CVIB can be reorganized as follow:

$$\frac{1}{|\mathcal{D}_B|} \sum_{(u,i)\in\mathcal{D}_B} l\left(h\left(u,i\right), f\left(u,i\right)\right) + \beta l\left(m_U, m_B\right). \quad (4)$$

For brevity, here we omit the irrelevant terms and only preserve the key modules. m_B and m_U are defined as follows:

$$\begin{aligned} m_B &= \frac{1}{|\mathcal{D}_B|} \sum_{(u,i)\in\mathcal{D}_B} h\left(g\left(u,i\right)\right), \\ m_U &= \frac{1}{|\mathcal{D}_U|} \sum_{(u,i)\in\mathcal{D}_U} h\left(g\left(u,i\right)\right), \end{aligned} \quad (5)$$

The second term is derived from information bottleneck, which reduces the discrepancy between the mean of the model predictions on biased training data and unbiased data. Note minimizing distribution discrepancy w.r.t. model's prediction boost representation invariance indirectly

DAMF [Saito and Nomura, 2022]. The objective can be rewritten as:

$$\frac{1}{|\mathcal{D}_B|} \sum_{(u,i)\in\mathcal{D}_B} l\left(h\left(u,i\right), f\left(u,i\right)\right) + \beta d_{h,\mathcal{H}}\left(\mathcal{D}_B, \mathcal{D}_U\right). \quad (6)$$

where $D_{h,\mathcal{H}}$ is a kind of metric of the distribution discrepancy. DAMF minimizes it between predictions on biased and unbiased data by adversarial learning.

InvPref [Wang *et al.*, 2022]. The objective of InvPref can be formulated as:

$$\frac{1}{|\mathcal{D}_B|} \sum_{(u,i)\in\mathcal{D}_B} l\left(h\left(u,i\right), f\left(u,i\right)\right) + \beta d_*\left(\mathcal{D}_1, \mathcal{D}_2, \cdots\right), \quad (7)$$

where d_* is defined by the min-max game in the original and measures the discrepancy among representations of multi-distributions. InvPref constructs heterogeneous environments $\{\mathcal{D}_1, \mathcal{D}_2, \cdots\}$ with different distributions via clustering. Based on the assumption that the distribution of unbiased data is a combination of distributions of the constructed environments, this strategy would naturally reduce representation distribution discrepancy between biased and unbiased data.

Based on the above analysis, the merit of existing IR-based methods can be easily understood — they employ various forms of regularizer to boost representation invariance, and reduce $\varepsilon_U(h, f)$ to some extent, yielding better performance.

Limitation of IR-based methods. Despite promising, we argue one limitation of existing invariant representation-based methods — **the representation discriminability has been overlooked.** These methods typically assume that the representation discriminability is sufficiently high, which however is not true in RS. In fact, RS usually suffers from extreme data sparsity and severe label shift. The increase in data sparsity will incur inadequate training of the model and inevitably hurts the representation discriminability. Worse still, purely pursuing the invariance will ruin the discriminability as the presence of label shift. This is because the optimum will force the representation of some user-item pairs with different labels to be close. Thus, towards better unbiased recommendation, it is essential to consider both representation invariance and discriminability.

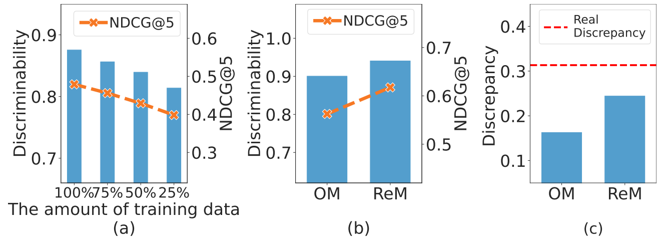

Figure 2: Discriminbility analysis in Yahoo!R3. (a) Impact of data volume on representation discriminability and recommendation performance by MF. (b) Impact of label shift on representation discriminability and recommendation performance by adversarial distribution alignment. OM and ReM refer to the learned invariant model with shift and without label shift respectively. (c) Comparison of the distribution discrepancy between predictions on biased and unbiased data and real label distribution discrepancy between biased and unbiased data.

2.3 Empirical Analysis

In this section, we conduct the empirical analysis on a real-world dataset Yahoo!R3 to provide evidence of how data sparsity and label shift hurt representation discriminability.

Data sparsity. We respectively use 100%, 75%, 50%, and 25% of biased data in Yahoo!R3 for training the Matrix Factorization (MF) model. We use $1 - \lambda^*$ [Chen *et al.*, 2019; Kundu *et al.*, 2022] to measure representation discriminability. The discriminability with varying data ratios of biased data is presented in Figure 2. As can be seen, the discriminability and performance drop heavily with the decrease in the data volume (*i.e.,* the increase of the data sparsity).

Label shift. To investigate the negative effect of label shift on discriminability, we conduct a comparable experiment. We create a new dataset from the original Yahoo!R3, where the biased data has been resampled (*i.e.,* under-sampling the positive instances) so that the resampled share the same label distribution as the unbiased. We then train two models (named OM and ReM respectively) on the biased data and the resampled respectively based on the common invariance boosting method[2] [Ganin and Lempitsky, 2015]. Additionally, we explore the distribution of predictions by the model

[2]The difference between the two models is only on different datasets involved in the calculation of the distribution discrepancy to the unbiased data, and the empirical errors are both calculated on the biased data.

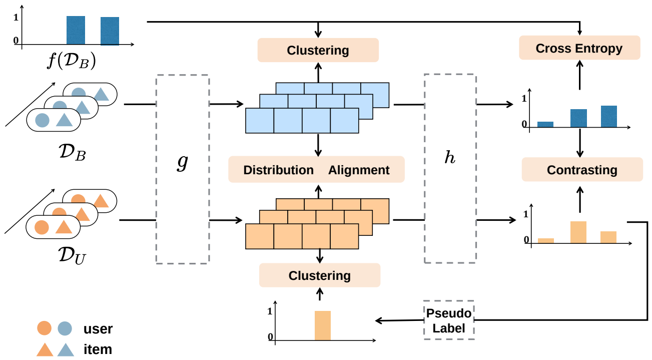

Figure 3: The overall framework of DIRL.

and study the discrepancy between the biased data and unbiased data. The results are presented in Figure 2. We make two interesting observations: 1) The discriminability and performance on the ReM are significantly larger than OM, suggesting label shift indeed hurts representation discriminability and recommendation performance; 2) In comparison to the distribution discrepancy between the predictions of ReM on biased and unbiased data, the distribution discrepancy between the predictions of OM on biased and unbiased data is significantly lower than the real label distribution discrepancy. This confirms that the blindly boosting representation invariance will force the representation of some user-item pairs with different labels to be close.

3 Proposed Method: DIRL

We now present the proposed DIRL framework (*cf.* Figure 3), which consists of three modules: 1) Adversarial distribution alignment for boosting representation invariance; 2) Label-conditional clustering for mitigating the ruin of data sparsity on representation discriminability; 3) Prior-guided contrasting for mitigating the ruin of label shift on discriminability.

3.1 Adversarial Distribution Alignment

Inspired by the success of adversarial training on UDA task [Jiang *et al.*, 2022], here we leverage adversarial learning in RS for mitigating the distribution discrepancy. That is, the metric of distribution discrepancy $d_{\mathcal{H}\Delta\mathcal{H}}(\mathcal{D}_B, \mathcal{D}_U)$ can be approximated by training a distribution classifier C.

Formally, the recommendation model and the classifier play a min-max game with optimizing:

$$\begin{aligned}\min_{\theta}\max_{\phi} L_d = &-\sum_{(u,i)\in\mathcal{D}_B} l\left(C\left(g\left(u,i\right)\right),1\right)\\ &-\sum_{(u,i)\in\mathcal{D}_U} l\left(C\left(g\left(u,i\right)\right),0\right).\end{aligned} \tag{8}$$

where θ and ϕ are the recommendation model's and the distribution classifier's parameters respectively. We use MF as our recommendation model. This means $g(u,i) := \mathbf{e}_u \odot \mathbf{e}_i$, where $\mathbf{e}_u$ and $\mathbf{e}_i$ are embeddings of the user u and the item i respectively. And $h(\mathbf{z}) := \sigma(\mathbf{z}^{\intercal}\mathbf{1})$, which is fixed as $\mathbf{1}$ in $\mathcal{H}$. In practice, the issue of serious gradient vanishing is encountered [Tzeng *et al.*, 2017] and thus a gradient trick [Tzeng *et al.*, 2015] has been utilized, *i.e.,* optimizing the recommendation model with:

$$\min_{\theta} L_{ada} = \sum_{(u,i)\in\{\mathcal{D}_B,\mathcal{D}_U\}} l\left(C\left(g\left(u,i\right)\right),\frac{1}{2}\right). \tag{9}$$

3.2 Prior-guided Contrasting

To mitigate the impact of label shift on representation discriminability, a straightforward solution is to directly restrict the distribution discrepancy of model predictions. Here we propose prior-guided contrasting that constrains the distance between the predictions on the biased data and unbiased data according to the prior knowledge of label shift. In a typical RS, observed training data usually exhibit much more ratio of positive instances than the test data collected from unbiased random policy. Naturally, a constraint is introduced for capturing such useful prior knowledge:

$$\begin{aligned} L_{con} &= -log\left(\sigma\left(\bar{m}_B - \bar{m}_U\right)\right),\\ \bar{m}_B &= \frac{1}{|\mathcal{D}_B|}\sum_{(u,i)\in\mathcal{D}_B}\bar{h}\left(g\left(u,i\right)\right),\\ \bar{m}_U &= \frac{1}{|\mathcal{D}_U|}\sum_{(u,i)\in\mathcal{D}_U}\bar{h}\left(g\left(u,i\right)\right).\end{aligned} \tag{10}$$

where $\bar{h}(.,.)$ denotes the predicted logits of the model. Here we directly constrain the average logit of training data shall be larger than test data.

Discussion. To better understand the effect, we conduct analyses based on the gradient *w.r.t.* model parameters θ:

$$-log\left(\sigma\left(1-\sigma\left(\bar{m}_B-\bar{m}_U\right)\right)\right)\frac{\partial\left(\bar{m}_B-\bar{m}_U\right)}{\partial\theta}. \tag{11}$$

As can be seen, each gradient step has a multiplicative scalar:

$$\Delta_{\bar{m}_B,\bar{m}_U} = -log\left(\sigma\left(1-\sigma\left(\bar{m}_B-\bar{m}_U\right)\right)\right). \tag{12}$$

This quantity depends on the gap of the average logit (*i.e.,* $\bar{m}_B - \bar{m}_U$). When it is contrary to the prior knowledge of real label distribution discrepancy, $\Delta_{\bar{m}_B,\bar{m}_U}$ is large, which makes a large update to the representation. When it is consistent with the prior, $\Delta_{\bar{m}_B,\bar{m}_U}$ will gradually decrease with the increase of the gap, which can prevent the gap from being too large. Minimizing Eq. 11 causes invariance boosting focus on eliminating distribution discrepancy outside of label shift, which prevents it from forcing the representation of user-item pairs with different labels to be close.

3.3 Label-conditional Clustering

Label-conditional clustering has been leveraged for boosting representation discriminability. A clustering loss is introduced, which can encourage representations with the same labels to gather closely together and those with different labels to be separated farther apart. Specifically, we optimize the following loss on biased data:

$$\begin{aligned} L_{cb} = &\frac{1}{|\mathcal{D}_B|}\sum_{y\in\{0,1\}}\sum_{(u,i)\in\mathcal{D}_B}\delta\left(f\left(u,i\right)=y\right)\left\|g\left(u,i\right)-c_B^y\right\|^2\\ &-\left\|c_B^{y=1}-c_B^{y=0}\right\|^2,\end{aligned} \tag{13}$$

Algorithm 1 DIRL

Input: History feedback D_B and $f(D_B)$; the user set $\mathcal{U}$ and the item set $\mathcal{I}$; trade-off parameters β, α_1, and α_2; learning rate lr; weight decay λ

Parameter: The recommendation model's parameter θ; the distribution classifier's parameter ϕ

Output: Model predictions of users' feedback on items

```
Randomly initialize θ and φ
while not convergence do
    Construct unbiased user-item pairs D_U by uniformly
    sampling n users and n items in U and I respectively
    Update φ by Adam according to max-step in Eq. 16
    Update θ by Adam according to min-step in Eq. 16
end while
return Recommendation model
```

where δ is the indicator function and c_B^y is the centroid of the representations with the same labels in biased data. c_B^y is defined as:

$$c_B^y = \frac{\sum_{(u,i)\in\mathcal{D}_B} \delta\left(f\left(u,i\right)=y\right) g\left(u,i\right)}{\sum_{(u,i)\in\mathcal{D}_B} \delta\left(f\left(u,i\right)=y\right)}. \tag{14}$$

We also apply the clustering loss on unbiased data to further enhance representation discriminability. Considering the labels of unbiased data are unavailable, inspired by the self-labeling technique [Lee and others, 2013], we use pseudo labels from model prediction as substitutes. That is, we simply assign the pseudo-label as 1 if the model prediction is greater than threshold $t_p = 0.5$, or to 0 if it is smaller. The overall cluster loss on both biased and unbiased data is:

$$L_{clu} = L_{cb} + L_{cu}. \tag{15}$$

Minimizing Eq. (15) increases the distance among representation centroids of different labels. It naturally mitigate the impact caused by data sparsity that representations with different labels are not separated enough.

3.4 Joint Optimization

DIRL pursues both representation invariance and discriminability, and optimizes the following joint objective function:

$$\begin{aligned} &\max_{\phi} L_d, \\ &\min_{\theta} \hat{\varepsilon}_B(h,f) + \beta L_{ada} + \alpha_1 L_{clu} + \alpha_2 L_{con}. \end{aligned} \tag{16}$$

where β, α_1 and α_2 are hyper-parameters that regulate the effects of the modules.

4 Experiments

In this section, we conduct several experiments to answer the following questions: **RQ1** Can DIRL outperform existing recommendation methods for mitigating selection bias? **RQ2** To what extent different components contribute to the effectiveness of DIRL? **RQ3** Does DIRL improve the representation discriminability and reduce the empirical error on the unbiased data as expected?

Table 1: The statistics of Yahoo!R3, KuaiRand-Pure, and Coat.

Dataset	#User	#Item	#Biased Data	#Unbiased Data
Yahoo!R3	15,400	1,000	311,704	5,400
KuaiRand-Pure	27,077	7,551	1,375,000	1,177,026
Coat	290	300	6,960	4,640

4.1 Experimental Setup

Datasets. We use three publicly available datasets: Yahoo!R3[3], Coat[4], and KuaiRand-Pure[5], which contain both biased data for training and unbiased data for testing. Following previous work [Chen *et al.*, 2021], we transform the ratings in Yahoo!R3 and Coat into positive (> 3) and negative (≤ 3) labels. We treat click or not in KuanRand-Pure as positive and negative labels. The statics of datasets are shown in Table 1.

Baselines. We compare DIRL with the following methods.

- **IPS** [Schnabel *et al.*, 2016]: IPS reweights the training data with item popularity as the propensity score.
- **DRJL** [Wang *et al.*, 2019]: DRJL performs propensity-based imputation learning, which combines IPS and the imputation model trained by IPS to train the base model.
- **MRDR** [Guo *et al.*, 2021]: MRDR extends DRJL by reducing the variance of propensity-based imputation learning.
- **CVIB** [Wang *et al.*, 2020]: CVIB employs an information contrastive loss and a prediction confidence penalty to balance the biased and unbiased data.
- **DAMF** [Saito and Nomura, 2022]. DAMF aligns the model prediction distributions of biased and unbiased data by adversarial learning.
- **InvPref** [Wang *et al.*, 2022]: Invpref iteratively construct heterogeneous environments by clustering in the training data and disentangles variant and invariant representation by adversarial learning.

All compared methods are based on the classic MF model. The hyperparameter settings of these methods are detailed in the Supplementary Material.

4.2 Performance Comparison (RQ1)

We first compare the ranking performance of DIRL with baselines. Table 2 shows results on Yahoo!R3, KuaiRand-Pure, and Coat. From the table, we observe that:

- In most cases, DIRL outperforms all baselines, with significant improvements, *e.g.,* the improvement on Yahoo!R3 is up to 9.25% *w.r.t.* NDCG@5. This result validates the effectiveness of DIRL and the rationality of optimizing both representation invariance and discriminability, which is consistant with our theoretical analysis.
- Propensity-based methods, IPS, DRJL, and MRDR, perform better than the base model MF, showing the effect of data reweighting. Furthermore, MRDR performs better than IPS and DRJL owing to its design for reducing variance in imputation learning. These results are consistent with previous works [Guo *et al.*, 2021].

[3]http://webscope.sandbox.yahoo.com/

[4]https://www.cs.cornell.edu/˜schnabts/mnar/

[5]https://kuairand.com/

Table 2: Recommendation performance. The bold and underlined fonts indicate the best and the second-best performance.

Method	Yahoo!R3			KuaiRand-Pure			Coat		
	NDCG@5	Precision@5	Recall@5	NDCG@5	Precision@5	Recall@5	NDCG@5	Precision@5	Recall@5
MF	0.4790	0.2318	0.6244	0.3271	0.2601	0.2704	0.5219	0.3143	0.5725
IPS	0.5048	0.2386	0.6446	0.3221	0.2556	0.2727	0.5429	0.3322	0.5979
DRJL	0.5503	0.2577	0.7139	0.3388	0.2687	0.2790	0.5515	0.3425	0.6236
MRDR	0.5686	0.2623	0.7276	0.3410	0.2693	0.2827	0.5593	0.3478	0.6261
CVIB	0.5356	0.2545	0.7048	0.3612	0.2799	0.3099	0.5645	0.3377	0.6141
DAMF	0.5679	0.2600	0.7169	0.3677	0.2860	0.3170	0.5693	0.3372	0.6108
InvPref	0.6229	0.2817	0.7723	0.3678	0.2868	0.3245	0.5926	0.3504	0.6488
DIRL w/o PL	0.5627	0.2599	0.7209	0.3508	0.2745	0.3027	0.5841	0.3432	0.6267
DIRL w/o P	0.6510	0.2873	0.7942	0.3784	0.2928	0.3227	0.6036	0.3497	0.6491
DIRL w/o L	0.6291	0.2808	0.7706	0.3872	0.2987	0.3308	0.5956	0.3451	0.6369
DIRL	**0.6805**	**0.2965**	**0.8151**	**0.3991**	**0.3063**	**0.3382**	**0.6101**	**0.3522**	**0.6536**
Impv (%)	9.25%	5.25%	5.54%	8.53%	6.84%	4.22%	2.95%	0.51%	0.73%

- In most cases, invariant representation-based methods especially InvPref outperforms propensity-based methods, validating the advantages of pursuing invariant representation to minimize the distribution discrepancy for addressing the selection bias in recommendation. Besides, this result indicates the potential of invariant representation learning for unbiased recommendation.
- DAMF performs better than CVIB. We postulate that the metric $d_{h,\mathcal{H}}$ utilized by DAMF can better measure the distribution discrepancy in recommendation than the mean-based measurement of CVIB.

4.3 Ablation Study (RQ2)

We then examine the impact of adversarial distribution alignment (A), prior-guided contrasting (P), and label-conditional clustering (L) on the effectiveness of DIRL by further comparing three variations of DIRL: (1) removal of both label-conditional clustering and prior-guided contrasting (DIRL w/o PL), (2) removal of prior-guided contrasting (DIRLw/o P), (3) removal of label-conditional clustering (DIRLw/o L), with DIRL and MF (*i.e.,* removal of adversarial distribution alignment, label-conditional clustering, and prior-guided contrasting). The results in Table 2 demonstrate that each module is critical to unbiased recommendations. In particular, (1) even ignoring representation discriminability (*i.e.,* DIRL w/o PL), it can achieve competitive performance, which shows the effectiveness of optimizing representation invariance through adversarial distribution alignment; (2) either label-conditional clustering (*i.e.,* DIRL w/o P) or prior-guided contrasting (*i.e.,* DIRL w/o L) improve recommendation performance effectively, showing the importance of optimizing representation discriminability; and (3) the best performance is achieved by using label-conditional clustering and prior-guided contrasting together (*i.e.,* DIRL), which further validates the effectiveness of jointly considering data sparsity and label shift.

We further evaluate the influence of adversarial distribution alignment, label-conditional clustering, and prior-guided contrasting on model performance by adjusting their coefficients (*i.e.,* β, α_1, and α_2). Figure 4 presents the results of model performance *w.r.t.* NDCG@5 as β, α_1, and α_2 varying in [0.0001, 0.001, 0.01, 0.1, 1.0, 10.0] on three datasets, respectively. Note that the performance *w.r.t.* Precision@5 and Recall@5 show similar trends (see Supplementary Material). From Figure 4, we find that: (1) DIRL can achieve optimal performance by striking a balance between different modules, which confirms the significance of promoting both representation invariance and discriminability; (2) the optimal values of α_1 and α_2 vary, as their corresponding modules boost representation discriminability by addressing issues arising from different sources (*i.e.,* label shift and data sparsity respectively); (3) unsuitable strength of each module deteriorates model performance.

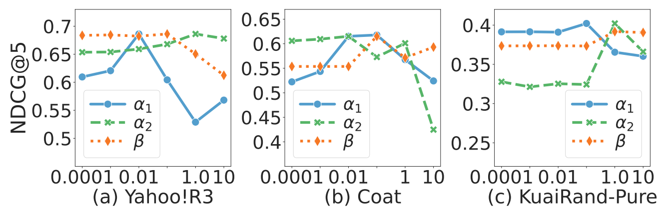

Figure 4: Hyperparameter sensitive analysis for NDCG@5 in Yahoo!R3, Coat, and KuaiRand-Pure.

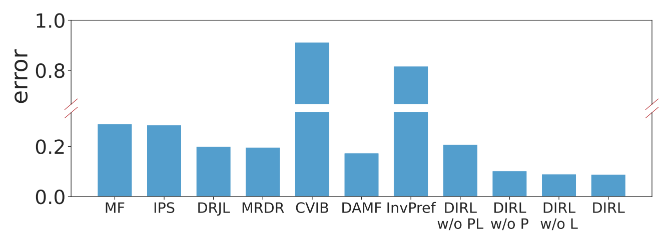

Figure 5: Empirical error on unbiased data in Yahoo!R3.

4.4 In-depth Analysis (RQ3)

We then investigate the sources of DIRL's performance gain by analyzing the representation.

Error on unbiased data. In addition to the superior ranking performance in Table 2, we further calculate the classification error of DIRL's representation on the unbiased data. Figure 5 shows that both prior-guided contrasting and label-conditional clustering reduces the error on unbiased data to varying degrees compared to other methods, which is consis-

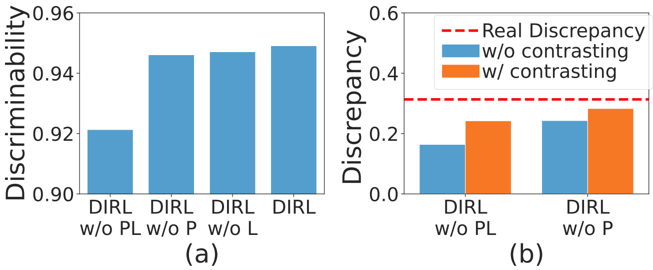

Figure 6: Discriminability analysis for DIRL in Yahoo!R3. (a) Discriminability of DIRL and its variants. (b) The distribution discrepancy between predictions on biased data and unbiased data.

tent with our conclusion in Section 2. As to the abnormal errors of CVIB and InvPref, we postulate the reason is that we tune hyperparameters by ranking performance but a high ranking performance may not obtain a low error.

Discriminability. We then evaluate the discriminability of DIRL representation. Figure 6 (a) shows that both prior-guided contrasting and label-conditional clustering improve the representation discriminability. This result confirms that improving representation discriminability indeed benefits the recommendation performance. Figure 6 (b) shows that both the DIRL w/o PL and DIRL w/o P with prior-guided contrasting have closer prediction distribution discrepancy to real label distribution discrepancy than without it. This means the prior-guided contrasting module improves discriminability in the way expected, *i.e.,* maintaining the true prediction distribution discrepancy between biased and unbiased data.

.

5 Related Work

In this section, we elaborate on the review of existing unbiased recommendation methods and unsupervised domain adaptation relating to our proposal.

Unbiased Recommendation. We focus on methods dealing with selection bias in recommendation. One of the most popular methods is to bridge the gap between biased and unbiased data. Propensity-based methods [Schnabel *et al.*, 2016; Imbens and Rubin, 2015] reweight biased data to align the distribution of unbiased data by the inverse propensity score. Various strategies have been proposed for estimating these propensity scores, such as using statistic metrics of users or items [Saito, 2020a], naive Bayes [Schnabel *et al.*, 2016], or logistic regression [Guo *et al.*, 2021]. [Saito, 2020b; Lee *et al.*, 2022] extend propensity-based methods from explicit feedback data to implicit feedback data. Additionally, methods such as doubly robust [Jiang and Li, 2016; Wang *et al.*, 2019; Liu *et al.*, 2020; Chen *et al.*, 2021; Dai *et al.*, 2022] and multiple robust [Li *et al.*, 2022] incorporate imputation learning to achieve double or multiple robustness for unbiased recommendation. However, propensity scores may extremely small and these propensity-based methods may have infinite bias, variance, and generalization error bounds [Guo *et al.*, 2021; Wu *et al.*, 2022]. An additional line for unbiased recommendation is to bridge the gap at the representation level [Liu *et al.*, 2020; Wang *et al.*, 2020; Saito and Nomura, 2022; Wang *et al.*, 2022]. [Wang *et al.*, 2020] leverage information bottleneck and obtain a contrastive loss for balancing the model's average predictions on biased and unbiased data. [Saito and Nomura, 2022] introduce UDA for unbiased recommendation and minimize distribution discrepancy between biased and unbiased data by adversarial distribution alignment. The current leading method is InvPref [Wang *et al.*, 2022], which disentangles variant and invariant representations by clustering and aligning distributions. However, these methods neglect the discriminability of representation.

Unsupervised Domain Adaptation. The task of UDA [Ben-David *et al.*, 2010] is to transfer knowledge from the labeled source domain to the unlabeled target domain. Adversarial distribution alignment is the most popular method in UDA, which focuses on minimizing the distribution discrepancy *w.r.t.* representation between source and target domains by adversarial learning [Ganin and Lempitsky, 2015; Tzeng *et al.*, 2015; Tzeng *et al.*, 2017; Zhang *et al.*, 2019; Acuna *et al.*, 2021]. However, these methods ignore the representation discriminability, leading to limited generalization performance [Chen *et al.*, 2019]. Some works balance representation invariance and discriminability for better generalization [Chen *et al.*, 2019; Kundu *et al.*, 2022]. Nevertheless, these works are not tailored to recommendation, lacking the consideration of label shift and data sparsity in recommendation. Along this line, a series of methods dealing with the mixed case of covariate distribution shift and label shift [Yan *et al.*, 2017; Deng *et al.*, 2019; Kang *et al.*, 2019; Wu *et al.*, 2019; Li *et al.*, 2020; Prabhu *et al.*, 2021; Liu *et al.*, 2021]. In an orthogonal direction, we propose a simple and effective module to reduce the impact of label shift on representation discriminability with prior knowledge of label distribution discrepancy. Note that such prior knowledge is unavailable in other scenarios.

6 Conclusion

In this paper, we studied the selection bias in recommender systems from the perspective of invariant representation. According to the analysis of the theory in UDA, we pointed out the importance of optimizing both the invariance and discriminability of representation. Furthermore, we empirically found that the presence of data sparsity and label shift can hurt the discriminability of representation. Accordingly, we proposed a new Discriminative-Invariant Representation Learning framework for unbiased recommendation with additional modules for counteracting the impact of data sparsity and label shift.

One interesting direction for future work is extending our methods to implicit feedback data, which can be collected easier. Besides, it is valuable to explore unbiased methods for sequential recommendation and conversation recommendation to further account for dynamic user preference.

References

[Acuna *et al.*, 2021] David Acuna, Guojun Zhang, Marc T Law, and Sanja Fidler. f-domain adversarial learning: Theory and algorithms. In *International Conference on Machine Learning*, pages 66–75. PMLR, 2021.

[Ben-David *et al.*, 2010] Shai Ben-David, John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, and Jennifer Wortman Vaughan. A theory of learning from different domains. *Machine learning*, 79(1):151–175, 2010.

[Chen *et al.*, 2019] Xinyang Chen, Sinan Wang, Mingsheng Long, and Jianmin Wang. Transferability vs. discriminability: Batch spectral penalization for adversarial domain adaptation. In *International conference on machine learning*, pages 1081–1090. PMLR, 2019.

[Chen *et al.*, 2020] Jiawei Chen, Hande Dong, Xiang Wang, Fuli Feng, Meng Wang, and Xiangnan He. Bias and debias in recommender system: A survey and future directions. *arXiv preprint arXiv:2010.03240*, 2020.

[Chen *et al.*, 2021] Jiawei Chen, Hande Dong, Yang Qiu, Xiangnan He, Xin Xin, Liang Chen, Guli Lin, and Keping Yang. Autodebias: Learning to debias for recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 21–30, 2021.

[Dai *et al.*, 2022] Quanyu Dai, Haoxuan Li, Peng Wu, Zhenhua Dong, Xiao-Hua Zhou, Rui Zhang, Rui Zhang, and Jie Sun. A generalized doubly robust learning framework for debiasing post-click conversion rate prediction. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 252–262, 2022.

[Deng *et al.*, 2019] Zhijie Deng, Yucen Luo, and Jun Zhu. Cluster alignment with a teacher for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9944–9953, 2019.

[Ganin and Lempitsky, 2015] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pages 1180–1189. PMLR, 2015.

[Guo *et al.*, 2021] Siyuan Guo, Lixin Zou, Yiding Liu, Wenwen Ye, Suqi Cheng, Shuaiqiang Wang, Hechang Chen, Dawei Yin, and Yi Chang. Enhanced doubly robust learning for debiasing post-click conversion rate estimation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 275–284, 2021.

[Huang *et al.*, 2022] Jin Huang, Harrie Oosterhuis, and Maarten de Rijke. It is different when items are older: Debiasing recommendations when selection bias and user preferences are dynamic. In *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, pages 381–389, 2022.

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

[Jiang and Li, 2016] Nan Jiang and Lihong Li. Doubly robust off-policy value evaluation for reinforcement learning. In *International Conference on Machine Learning*, pages 652–661. PMLR, 2016.

[Jiang *et al.*, 2022] Junguang Jiang, Yang Shu, Jianmin Wang, and Mingsheng Long. Transferability in deep learning: A survey. *CoRR*, abs/2201.05867, 2022.

[Kang *et al.*, 2019] Guoliang Kang, Lu Jiang, Yi Yang, and Alexander G Hauptmann. Contrastive adaptation network for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4893–4902, 2019.

[Koren *et al.*, 2009] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

[Kundu *et al.*, 2022] Jogendra Nath Kundu, Akshay R Kulkarni, Suvaansh Bhambri, Deepesh Mehta, Shreyas Anand Kulkarni, Varun Jampani, and Venkatesh Babu Radhakrishnan. Balancing discriminability and transferability for source-free domain adaptation. In *International Conference on Machine Learning*, pages 11710–11728. PMLR, 2022.

[Lee and others, 2013] Dong-Hyun Lee et al. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In *Workshop on challenges in representation learning, ICML*, volume 3, page 896, 2013.

[Lee *et al.*, 2022] Jae-woong Lee, Seongmin Park, Joonseok Lee, and Jongwuk Lee. Bilateral self-unbiased learning from biased implicit feedback. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '22, page 29–39, New York, NY, USA, 2022. Association for Computing Machinery.

[Li *et al.*, 2020] Bo Li, Yezhen Wang, Tong Che, Shanghang Zhang, Sicheng Zhao, Pengfei Xu, Wei Zhou, Yoshua Bengio, and Kurt Keutzer. Rethinking distributional matching based domain adaptation. *CoRR*, abs/2006.13352, 2020.

[Li *et al.*, 2022] Haoxuan Li, Quanyu Dai, Yuru Li, Yan Lyu, Zhenhua Dong, Peng Wu, and Xiao-Hua Zhou. Multiple robust learning for recommendation. *arXiv preprint arXiv:2207.10796*, 2022.

[Liu *et al.*, 2020] Dugang Liu, Pengxiang Cheng, Zhenhua Dong, Xiuqiang He, Weike Pan, and Zhong Ming. A general knowledge distillation framework for counterfactual recommendation via uniform data. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 831–840, 2020.

[Liu *et al.*, 2021] Xiaofeng Liu, Zhenhua Guo, Site Li, Fangxu Xing, Jane You, C-C Jay Kuo, Georges El Fakhri, and Jonghye Woo. Adversarial unsupervised domain adaptation with conditional and label shift: Infer, align and iterate. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10367–10376, 2021.

[Prabhu *et al.*, 2021] Viraj Prabhu, Shivam Khare, Deeksha Kartik, and Judy Hoffman. Sentry: Selective entropy optimization via committee consistency for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Interna-

tional Conference on Computer Vision, pages 8558–8567, 2021.

[Saito and Nomura, 2022] Yuta Saito and Masahiro Nomura. Towards resolving propensity contradiction in offline recommender learning. In Lud De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 2211–2217. International Joint Conferences on Artificial Intelligence Organization, 7 2022. Main Track.

[Saito, 2020a] Yuta Saito. Asymmetric tri-training for debiasing missing-not-at-random explicit feedback. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 309–318, 2020.

[Saito, 2020b] Yuta Saito. Unbiased pairwise learning from biased implicit feedback. In *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*, pages 5–12, 2020.

[Schnabel *et al.*, 2016] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. Recommendations as treatments: Debiasing learning and evaluation. In *international conference on machine learning*, pages 1670–1679. PMLR, 2016.

[Swaminathan and Joachims, 2015] Adith Swaminathan and Thorsten Joachims. The self-normalized estimator for counterfactual learning. *advances in neural information processing systems*, 28, 2015.

[Tzeng *et al.*, 2015] Eric Tzeng, Judy Hoffman, Trevor Darrell, and Kate Saenko. Simultaneous deep transfer across domains and tasks. In *Proceedings of the IEEE international conference on computer vision*, pages 4068–4076, 2015.

[Tzeng *et al.*, 2017] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7167–7176, 2017.

[Wang *et al.*, 2019] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. Doubly robust joint learning for recommendation on data missing not at random. In *International Conference on Machine Learning*, pages 6638–6647. PMLR, 2019.

[Wang *et al.*, 2020] Zifeng Wang, Xi Chen, Rui Wen, Shao-Lun Huang, Ercan Kuruoglu, and Yefeng Zheng. Information theoretic counterfactual learning from missing-not-at-random feedback. *Advances in Neural Information Processing Systems*, 33:1854–1864, 2020.

[Wang *et al.*, 2022] Zimu Wang, Yue He, Jiashuo Liu, Wenchao Zou, Philip S Yu, and Peng Cui. Invariant preference learning for general debiasing in recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1969–1978, 2022.

[Wu *et al.*, 2019] Yifan Wu, Ezra Winston, Divyansh Kaushik, and Zachary Lipton. Domain adaptation with asymmetrically-relaxed distribution alignment. In *International conference on machine learning*, pages 6872–6881. PMLR, 2019.

[Wu *et al.*, 2022] Peng Wu, Haoxuan Li, Yan Lyu, Chunyuan Zheng, and Xiao-Hua Zhou. Doubly robust collaborative targeted learning for debiased recommendations. *arXiv preprint arXiv:2203.10258*, 2022.

[Yan *et al.*, 2017] Hongliang Yan, Yukang Ding, Peihua Li, Qilong Wang, Yong Xu, and Wangmeng Zuo. Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2272–2281, 2017.

[Zhang *et al.*, 2019] Yuchen Zhang, Tianle Liu, Mingsheng Long, and Michael Jordan. Bridging theory and algorithm for domain adaptation. In *International Conference on Machine Learning*, pages 7404–7413. PMLR, 2019.

[Zhao *et al.*, 2019] Han Zhao, Remi Tachet Des Combes, Kun Zhang, and Geoffrey Gordon. On learning invariant representations for domain adaptation. In *International Conference on Machine Learning*, pages 7523–7532. PMLR, 2019.