UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback

Gupta, S.; Oosterhuis, H.; de Rijke, M.

DOI
10.1145/3578337.3605114

Publication date
2023

Document Version
Final published version

Published in
ICTIR '23

License
CC BY

Link to publication

Citation for published version (APA):
Gupta, S., Oosterhuis, H., & de Rijke, M. (2023). A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback. In *ICTIR '23: Proceedings of the 2023 ACM SIGIR International Conference on Theory of Information Retrieval : July 23, 2023, Taipei, Taiwan* (pp. 87–93). Association for Computing Machinery. https://doi.org/10.1145/3578337.3605114

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Oct 2026

A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback

Shashank Gupta
University of Amsterdam
Amsterdam, The Netherlands
s.gupta2@uva.nl

Harrie Oosterhuis
Radboud Universiteit
Nijmegen, The Netherlands
harrie.oosterhuis@ru.nl

Maarten de Rijke
University of Amsterdam
Amsterdam, The Netherlands
m.derijke@uva.nl

ABSTRACT

Variational autoencoders (VAEs) are the state-of-the-art model for recommendation with implicit feedback signals. Unfortunately, implicit feedback suffers from selection bias, e.g., popularity bias, position bias, etc., and as a result, training from such signals produces biased recommendation models. Existing methods for debiasing the learning process have not been applied in a generative setting. We address this gap by introducing an inverse propensity scoring (IPS) based method for training VAEs from implicit feedback data in an unbiased way. Our IPS-based estimator for the VAE training objective, VAE-IPS, is provably unbiased w.r.t. selection bias. Our experimental results show that the proposed VAE-IPS model reaches significantly higher performance than existing baselines. Our contributions enable practitioners to combine state-of-the-art VAE recommendation techniques with the advantages of bias mitigation for implicit feedback.

CCS CONCEPTS

• **Information systems** → **Recommender systems**.

KEYWORDS

Variational autoencoder; Implicit feedback; Selection bias

ACM Reference Format:
Shashank Gupta, Harrie Oosterhuis, and Maarten de Rijke. 2023. A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback. In *Proceedings of the 2023 ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR '23), July 23, 2023, Taipei, Taiwan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3578337.3605114

1 INTRODUCTION

Recommender systems rely on user feedback signals to infer user preferences [14, 39]. User feedback can be classified into *explicit* and *implicit* signals. Ratings are a prime example of explicit feedback signals, which can be interpreted as the users' preference over items [29]. However, user ratings are sparse in nature, which makes learning from them difficult [2]. Implicit feedback signals, such as clicks, purchases, or views, are abundantly present in the form of interactions, enabling the learning of user preferences on a large scale [16, 17].

Despite the advantages of learning from implicit feedback data, such data suffers from various forms of bias, and consequently, it generally does not reflect the *true* user preferences but is a biased indicator of it [17, 18, 25, 28]. Examples include (i) popularity bias [2, 29, 37], where some items get more clicks due to their popularity on the platform; (ii) positivity bias [15, 29], where users are more likely to provide ratings for items they would rate highly; and (iii) position bias [5, 43], where users tend to click items ranked at a higher position of the result page. To mitigate the negative effects of selection bias, prior work has proposed the usage of inverse propensity scoring (IPS), a counterfactual estimation technique [35]. IPS counteracts the effect of selection bias by reweighting datapoints inversely to their probability of being observed. Recently, the IPS approach has been extended to optimize matrix factorization (MF) models from implicit feedback [34]. While MF methods [13, 14] and more recent neural MF-based methods [12, 46] have a long tradition in the recommendation field, the state-of-the-art methods for learning from implicit feedback data use variational autoencoders (VAEs) instead [21, 36, 41]. VAE-based methods perform well in the low-data regime, when only a few interactions are available for most users [6, 21]. Despite the importance of bias mitigation and the strong performance of VAEs for recommendation from implicit feedback, existing work has overlooked the issue of bias in a generative setting [20, 34, 35]. To the best of our knowledge, previous work has not considered state-of-the-art debiasing combined with VAEs, despite the obvious potential for improved performance.

We address this gap by introducing VAE-IPS: an IPS debiasing method for VAE optimization from implicit feedback. We start by introducing an ideal generative objective for training with implicit feedback data. We then show that naively ignoring selection bias during learning leads to the optimization of a biased estimate of the ideal generative objective. In contrast, we propose VAE-IPS, which uses IPS, and prove that it unbiasedly optimizes the ideal generative objective in expectation. Our experiments show that VAE-IPS reaches higher performance than existing debiasing methods and VAE without debiasing. Thus, VAE-IPS successfully integrates state-of-the-art VAE recommendation models with the advantages of debiasing implicit feedback.

2 RELATED WORK

Collaborative filtering (CF) is often approached as a matrix completion problem. A general problem with recommendation data is that it is missing-not-at-random [22]. Missing entries should not be treated as negative feedback because their absence is often not due to a lack of preference. Training standard CF-based methods on such biased data leads to suboptimal learning and evaluation [23, 24]. A common technique to learn and evaluate in an unbiased manner is based on IPS. It has been applied in multiple feedback settings:

This work is licensed under a Creative Commons Attribution International 4.0 License.

ICTIR '23, July 23, 2023, Taipei, Taiwan
© 2023 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0073-6/23/07.
https://doi.org/10.1145/3578337.3605114

(i) explicit CF [15, 35], (ii) pointwise implicit [34], and (iii) pairwise implicit [32]. In this work, we extend the application of debiasing methods to deep generative models.

Latent factor models are a popular choice for training recommender systems [3, 8, 30, 45], but they are limited to linear models. Alternatively, deep neural networks are also used to model recommendation problems [12, 21, 46]. In particular, VAE-based models have strong and robust CF performance [6, 21]. The *BiVAE* model models the *dyadic* nature of user-item interaction data [41]. BiVAE models the distribution of user-item scores as a Bernoulli distribution, instead of a multinomial distribution over all items for a given user [6, 21]. Our work builds on the BiVAE framework by contributing an unbiased generative model of user-item relevance.

3 BACKGROUND

Selection bias in click feedback. Let $r_{u,i}$ be an indicator variable for the *true* user-item relevance, and let $c_{u,i}$ be a binary click indicator variable, indicating whether user u clicked item i. Lastly, $o_{u,i}$ denotes the observation variable, indicating whether user u observed item i. We assume that interaction behavior follows a simple examination model, where the click probability is a product of the probability of observance and that of relevance [4, 27, 32, 34, 42]:

$$p(c_{u,i} = 1) = p(o_{u,i} = 1) \cdot p(r_{u,i} = 1) = \rho_{u,i} \cdot \gamma_{u,i}, \tag{1}$$

where $\rho_{u,i}$ and $\gamma_{u,i}$ are the observation and relevance probability, respectively. Intuitively, the click model implies that clicks only occur on items that are relevant to a user and observed by them.

Variational autoencoders for click feedback. The goal of a VAE is to model the generative distribution $p(c_{ui} \mid z_u, z_i)$ [41], where c_{ui} is the click signal, and z_u, z_i are k-dimensional latent variables for the corresponding user-item pair. Typically, we are interested in the posterior distribution $p(z_u, z_i \mid r_{ui})$, which can be used for Bayesian inference post learning. We follow the recently proposed VAE model *BiVAE*, which models the pointwise generative distribution of the interaction signal.

Next, we define the ideal generative objective for user-item relevance following the BiVAE [41] framework. The generative story for user-item relevance is defined as follows:

- for each user u in the dataset, draw a latent vector $z_u \sim \mathcal{N}(0, I_k)$;
- for each item i in the dataset, draw a latent vector $z_i \sim \mathcal{N}(0, I_k)$;
- for each user-item pair (u, i) in the dataset, the relevance score is defined as follows: $s_{u,i} = \sigma(z_u^T z_i)$, where the function σ is the sigmoid function ($\sigma(x) = \frac{1}{1+e^{-x}}$); and
- for each user-item pair (u, i), draw the relevance variable $r_{u,i}$ from a Bernoulli distribution as follows: $r_{u,i} \sim Bern(s_{u,i})$.

The log-likelihood for the relevance as defined previously (Eq. 1) can be expressed as:

$$L^{rel} = \frac{1}{N} \sum_{(u,i)} L^{rel}_{u,i}, \qquad L^{rel}_{u,i} = \log p(r_{u,i}), \tag{2}$$

where N is the total number of interactions in the dataset, and $L^{rel}_{u,i}$ is the log-likelihood for the single (u, i) pair. Henceforth, the equations are defined for a single (u, i) interaction. For defining the *ideal* relevance generative objective, let the posterior for the user-item latent variables (z_u, z_i) (which we denote as $z_{u,i}$ going forward) be $q_\phi(z_{u,i})$ (with ϕ as parameters of the posterior network), and let the conditional likelihood distribution for relevance be $p_\theta(r_{u,i} \mid z_{u,i})$ (with θ as parameters of the conditional likelihood's network). Then, the likelihood in Eq. 2 is defined as:

$$\begin{aligned} L^{rel}_{u,i} &= \log \int p_\theta(r_{u,i}, z_{u,i}) \mathrm{d}z_{u,i} \\ &= \log \int p_\theta(r_{u,i}, z_{u,i}) \frac{q_\phi(z_{u,i})}{q_\phi(z_{u,i})} \mathrm{d}z_{u,i} \\ &= \log \mathbb{E}_{q_\phi(z_{u,i})} \left[\frac{p_\theta(r_{u,i}|z_{u,i}) p(z_{u,i})}{q_\phi(z_{u,i})} \right]. \end{aligned} \tag{3}$$

Next, we define $L^{ideal}_{u,i}$ as the following lower-bound of the log-likelihood:

$$\begin{aligned} L^{rel}_{u,i} &\geq \mathbb{E}_{q_\phi(z_{u,i})} \left[\log \frac{p_\theta(r_{u,i}|z_{u,i}) p(z_{u,i})}{q_\phi(z_{u,i})} \right] \\ &= \underbrace{\mathbb{E}_{q_\phi(z_{u,i})} \left[\log p_\theta(r_{u,i}|z_{u,i}) \right]}_{\text{Conditional likelihood}} - \underbrace{D\Big(q_\phi(z_{u,i}) \,\|\, p(z_{u,i}) \Big)}_{\text{KLD-regularizer}} = L^{ideal}_{u,i}, \end{aligned} \tag{4}$$

where we define $L^{ideal}_{u,i}$ as this lower-bound of the log-likelihood, also known as the evidence lower bound objective (ELBO) in the autoencoder literature [19, 41]. This is the *ideal* distribution, since the ELBO is the quantity that is optimized in VAEs instead of the log-likelihood ($L^{rel}_{u,i}$) [19]. The first term is the conditional-likelihood, and the second term is the KL-divergence between the posterior and the prior ($p(z_{u,i})$), which acts as a regularizer. In general, the following inequality holds: $L^{rel}_{u,i} \geq L^{ideal}_{u,i}$. We assume that relevance is a binary random variable (Bernoulli distributed) for implicit feedback data ($r_{ui} \in [0, 1]$), plugging it into Eq. 4, we obtain:

$$p_\theta(r_{u,i}|z_{u,i}) = (\pi_\theta(z_{u,i}))^{r_{ui}} (1 - \pi_\theta(z_{u,i}))^{1-r_{u,i}}, \tag{5}$$

where $\pi_\theta(z_{u,i})$ is the probability of relevance for the (u, i) pair. Then

$$\begin{aligned} L^{ideal}_{u,i} = \mathbb{E}_{q_\phi(z_{u,i})} \big[r_{ui} \log(\pi_\theta(z_{u,i})) + (1 - r_{u,i}) \log(1 - \pi_\theta(z_{u,i})) \big] \\ - D\big(q_\phi(z_{u,i}) \| p(z_{u,i})\big). \end{aligned} \tag{6}$$

The first part of the objective is similar to the familiar cross-entropy loss used with binary MF [14, 31], where in the VAE case, the latent embeddings are sampled from a distribution, as opposed to being deterministic in case of MF. The second part of the loss is the KL-divergence between the posterior $q_\phi(z_{u,i})$ and a simple normal distribution based prior $p(z_{u,i})$, which acts as a regularizer during the training procedure [19, 41].

4 METHOD: VAE-IPS ESTIMATOR

In this section, we introduce the naive click-based estimator, followed by a discussion of its bias, and finally we introduce our proposed estimator, which is unbiased in expectation.

Naively taking the existing loss (Eq. 4) and replacing relevance with clicks, results in the following biased estimate of the loss function:

$$\begin{aligned} L^{click}_{u,i} = \mathbb{E}_{q_\phi} \big[c_{ui} \log(\pi_\theta(z_{u,i})) + (1 - c_{u,i}) \log(1 - \pi_\theta(z_{u,i})) \big] \\ - D\big(q_\phi(z_{u,i}) \| p(z_{u,i})\big). \end{aligned} \tag{7}$$

In other words, the click-based objective function $L^{click}_{u,i}$ is simply

the ideal objective function (Eq. 4), where relevance $r_{u,i}$ has been subsituted by the click signal $c_{u,i}$. To prove that this is a biased estimator of the ideal objective, we apply the click model assumption and derive the expected value of this estimator with respect to the observation variable $o_{u,i}$, i.e., we consider $\mathbb{E}_o[L_{u,i}^{click}]$:

$$\begin{aligned}
&\mathbb{E}_o[L_{u,i}^{click}] + D(q_\phi(z_{u,i})\|p(z_{u,i})) \\
&= \mathbb{E}_{o,q_\phi}\big[c_{ui}\log(\pi_\theta(z_{u,i})) + (1-c_{u,i})\log(1-\pi_\theta(z_{u,i}))\big] \\
&= \mathbb{E}_{q_\phi}\big[\mathbb{E}_o[c_{u,i}]\log(\pi_\theta(z_{u,i})) + (1-\mathbb{E}_o[c_{u,i}])\log(1-\pi_\theta(z_{u,i}))\big] \\
&= \mathbb{E}_{q_\phi}\big[(\rho_{u,i}r_{u,i})\log(\pi_\theta(z_{u,i})) + (1-\rho_{u,i}r_{u,i})\log(1-\pi_\theta(z_{u,i}))\big].
\end{aligned} \tag{8}$$

Clearly, this is a biased estimate of the ideal loss (Eq. 4), where the propensity-term $\rho_{u,i}$ is a confounding variable. We can express the exact bias by the following difference:

$$\begin{aligned}
&\mathbb{E}_o[L_{u,i}^{naive}] - L_{u,i}^{ideal} \\
&\quad= \mathbb{E}_{q_\phi}\big[(1-\rho_{u,i})r_{u,i}\ \log(\pi_\theta(z_{u,i})) + \\
&\qquad\qquad (1-\rho_{u,i})r_{u,i}\log(1-\pi_\theta(z_{u,i}))\big] \\
&\quad= \mathbb{E}_{q_\phi}\left[(\rho_{u,i}-1)r_{u,i}\log\left(\frac{\pi_\theta(z_{u,i})}{1-\pi_\theta(z_{u,i})}\right)\right].
\end{aligned} \tag{9}$$

From Eq. 9 it is clear that the click-based estimator will be unbiased only if $\rho_{u,i} = 1$, for all (u,i) pairs, which is clearly an unfeasible condition with the prevalence of selection bias in interaction data.

4.1 Proposed unbiased generative estimator

We propose an unbiased generative estimator, in a similar vain as existing IPS corrections for position bias [43], trust-bias [1], and popularity and positivity-bias [15]. The IPS-corrected unbiased estimate of the true generative objective is defined as follows:

$$\begin{aligned}
L_{u,i}^{ips} = \mathbb{E}_{q_\phi}\left[\frac{c_{u,i}}{\rho_{u,i}}\log(\pi_\theta(z_{u,i})) + \left(1-\frac{c_{u,i}}{\rho_{u,i}}\right)\log(1-(\pi_\theta(z_{u,i})))\right] \\
- D\big(q_\phi(z_{u,i})\|p(z_{u,i})\big).
\end{aligned} \tag{10}$$

This estimator is an unbiased estimate of the true relevance based objective (Eq. 4). To prove this, we derive the expected value of the estimator with respect to the observation variable:

$$\begin{aligned}
&\mathbb{E}_o\big[L_{u,i}^{ips}\big] + D\big(q_\phi(z_{u,i})\|p(z_{u,i})\big) \\
&\quad= \mathbb{E}_{o,q_\phi}\left[\frac{c_{ui}}{\rho_{u,i}}\log(\pi_\theta(z_{u,i})) + \left(1-\frac{c_{ui}}{\rho_{u,i}}\right)\log(1-\pi_\theta(z_{u,i}))\right] \\
&\quad= \mathbb{E}_{q_\phi}\left[\mathbb{E}_o\left[\frac{c_{ui}}{\rho_{u,i}}\right]\log(\pi_\theta(z_{u,i})) \right. \\
&\qquad\qquad \left. + \left(1-\mathbb{E}_o\left[\frac{c_{ui}}{\rho_{u,i}}\right]\right)\log(1-\pi_\theta(z_{u,i}))\right] \\
&\quad= \mathbb{E}_{q_\phi}\big[r_{u,i}\log(\pi_\theta(z_{u,i})) + (1-r_{u,i})\log(1-\pi_\theta(z_{u,i}))\big].
\end{aligned} \tag{11}$$

Thus, we prove that in expectation $L_{u,i}^{ips}$ is equal to the ideal relevance-based objective from Eq. 4, and therefore, our VAE-IPS estimator is provenly unbiased.

4.2 Variance of the novel VAE-IPS estimator

Theorem 4.1. *The variance of the estimator $L_{u,i}^{ips}$, given the propensity $\rho_{u,i}$, click $c_{u,i}$, and relevance $r_{u,i}$, can be computed as:*

$$\begin{aligned}
Var(L_{u,i}^{ips}) = r_{u,i}\left(\frac{1}{\rho_{u,i}} - 1\right)\mathbb{E}_{q_\phi}\big[(\log(\pi_\theta(z_{u,i})) - \\
\log(1-\pi_\theta(z_{u,i})))^2\big].
\end{aligned} \tag{12}$$

Proof. We need to estimate the following quantity:

$$Var(L_{u,i}^{ips}) = \mathbb{E}_{q_\phi}\Big[\underbrace{\mathbb{E}_o[(L_{u,i}^{ips})^2]}_{(1)} - \underbrace{\mathbb{E}_o[L_{u,i}^{ips}]^2}_{(2)}\Big]. \tag{13}$$

To keep our notation concise, we will use $\delta_{u,i}^1 = \log(\pi_\theta(z_{u,i}))$ and $\delta_{u,i}^0 = \log(1-\pi_\theta(z_{u,i}))$. We start by expanding the first part of Eq. 13:

$$\begin{aligned}
\mathbb{E}_o[(L_{u,i}^{ips})^2] &= \mathbb{E}_o\left[\left(\frac{c_{ui}}{\rho_{u,i}}\delta_{u,i}^1 + (1-\frac{c_{ui}}{\rho_{u,i}})\delta_{u,i}^0\right)^2\right] \\
&= \mathbb{E}_o\Big[\frac{c_{ui}}{\rho_{u,i}^2}(\delta_{u,i}^1)^2 + (1-\frac{c_{ui}}{\rho_{u,i}})^2(\delta_{u,i}^0)^2 \\
&\qquad + 2(1-\frac{c_{ui}}{\rho_{u,i}})\frac{c_{ui}}{\rho_{u,i}}\delta_{u,i}^0\delta_{u,i}^1\Big] \\
&= \frac{r_{u,i}}{\rho_{u,i}}(\delta_{u,i}^1)^2 + \left(1-2r_{u,i}+\frac{r_{u,i}}{\rho_{u,i}}\right)(\delta_{u,i}^0)^2 \\
&\qquad + 2\left(r_{u,i}-\frac{r_{u,i}}{\rho_{u,i}}\right)\delta_{u,i}^0\delta_{u,i}^1,
\end{aligned} \tag{14}$$

where in the second step, we make use of the identity $c_{u,i}^2 = c_{u,i}$, and going from the second to third step, we evaluate the expectation term inside. Next, we expand the second part of the Eq. 13 by resolving the inner expectation with the use of Eq. 11:

$$\begin{aligned}
\mathbb{E}_o[L_{u,i}^{ips}]^2 &= \left(r_{u,i}\delta_{u,i}^1 + (1-r_{u,i})\delta_{u,i}^0\right)^2 \\
&= r_{u,i}(\delta_{u,i}^1)^2 + (1-r_{u,i})(\delta_{u,i}^0)^2,
\end{aligned} \tag{15}$$

where we also make use of the identity $r_{u,i}^2 = r_{u,i}$, and $(1-r_{u,i})^2 = 1-r_{u,i}$. Substituting the first and the second part back to Eq. 13:

$$\begin{aligned}
Var(L_{u,i}^{ips}) &= \mathbb{E}_{q_\phi}\left[r_{u,i}\left(\frac{1}{\rho_{u,i}}-1\right)\left((\delta_{u,i}^1)^2 + (\delta_{u,i}^0)^2 - 2\delta_{u,i}^0\delta_{u,i}^1\right)\right] \\
&= \mathbb{E}_{q_\phi}\left[r_{u,i}\left(\frac{1}{\rho_{u,i}}-1\right)(\delta_{u,i}^1 - \delta_{u,i}^0)^2\right].
\end{aligned}$$

□

We see that the variance depends inversely on the propensity, which suggests that items with lower propensity will have a higher variance, and vice-versa. In practice, this is not a big issue since methods like propensity clipping can greatly reduce variance at the cost of a small amount of bias. Eq. 16 shows how we apply this technique in our experimental setup.

5 EXPERIMENTAL SETUP

We assess the performance of VAE-IPS for a relevance prediction task in semi-synthetic, real-world, and fully-synthetic setups.

5.1 Baselines and settings

Our comparison includes the following six methods: (i) **Binary Matrix Factorization (MF)**: We use the matrix factorization model for implicit feedback dataset from [14], where the squared loss is

replaced with the cross-entropy loss to account for the clicks being Bernoulli distributed. (ii) **Rel-MF**: The binary matrix factorization model trained with IPS weighted loss from [34]. (iii) **MD-DR**: A doubly-robust variant of the IPS matrix factorization model, which uses a control variate to reduce the variance of the IPS method [44]. (iv) **MF-Dual**: The dual unbiased matrix factorization model for implicit feedback data. To the best of our knowledge, it is the current state-of-the-art method for debiasing implicit feedback data [20]. (v) **VAE**: We use the BiVAE framework developed to model dyadic data [41], which is more suitable for pointwise predictions. This VAE baseline is optimized with the proposed coordinate descent style optimization method, where the posteriors for user and items are optimized alternately. The latent variables for this case are z_u and z_i, and the user-item prediction score is defined as the dot product between the user and item latent variable, $\pi_\theta(z_{u,i}) = z_u^T \cdot z_i$. This is similar to the prediction score defined in matrix factorization, the key difference being the use of neural networks and variational inference. And (vi) **VAE-IPS**: This is our proposed method, the BiVAE model optimized with the unbiased VAE-IPS objective (Eq. 10). Practically, with the alternative coordinate descent optimization, we use the IPS correction alternately for both user-based and item-based loss functions in the BiVAE framework.

Variance reduction. The IPS estimator is known to suffer from large variance [40], due to the use of the inverse of the probability score, which is unbounded. We apply propensity clipping [34, 38], which can greatly reduce variance while only introducing a small amount of bias. Formally, we define a clipped propensity as:

$$\rho_{u,i}^{clip} = \max(\rho_{u,i}, \tau), \tag{16}$$

where τ is a hyper-parameter, which controls the trade-off between bias and the variance. A small τ can result in high variance but little bias, whereas a high τ can lead to little variance but high bias.

Propensity estimation. To estimate the propensity of an item, we use its relative click frequency in the training dataset [34]. The intuition behind training click frequency is that an item is more likely to be exposed to a user if it has historically been clicked more and vice-versa. Formally, we define the propensities as:

$$\rho_{ui} = \rho_i = \frac{\sum_{u=1}^{U} c_{u,i}}{\sum_{u=1}^{U}\sum_{i=1}^{I} c_{u,i}}, \tag{17}$$

where we make the assumption that propensity scores are uniform across all users [34].

Implementation details. The source code to reproduce the findings from the paper is available at: https://github.com/shashankg7/VAE-IPS.

5.2 Semi-synthetic experimental setup

To assess the performance of our proposed method, we use the MovieLens-1M dataset [11]. The dataset consists of ~6K users and ~3,700 items, with 1 million explicit feedback ratings. To convert an explicit feedback dataset into an implicit feedback dataset, we consider all ratings with a value over 4 as positive interactions and rest of the interactions as unlabelled instances.

We follow the experimental setup from [31]. To evaluate the performance of the methods, we use 50% of the dataset as test set. To simulate an unbiased test set, we re-sample 30% data from the test set with a sampling probability as $1/p_i^\alpha$, where p_i^α is an item's normalized frequency in the training dataset, and α is used to control the selection bias in the test set. A value of $\alpha = 1$ ensures the least selection bias and other values simulate controlled randomization.

We use NDCG@5 and MAP@5 as evaluation metrics. For calculating the normalizing factor of the NDCG@5, we follow the advice from [7] and base it on the entire dataset. Our evaluation metrics follow the definitions of earlier work on Rel-MF [34].

5.3 Real-world dataset experimental setup

We also evaluate VAE-IPS on a real-world dataset, where the test set interactions are from a truly uniform-random policy. We use the Yahoo! R3 dataset [23], which consists of interactions from a music recommendation service. The randomized test set ensures that it is free from the selection bias present in the training set.

To get a validation set, we split the training set in both datasets in accordance to a 80%/20% randomized split. We use the validation dataset to tune the hyper-parameters for the baselines and VAE-IPS. We use DCG@5 as the metric for hyper-parameter tuning, and tune the hyper-parameters using the self-normalized importance sampling (SNIPS) version of the DCG@5 metric [35].

5.4 Fully-synthetic experimental setup

We also examine a more controlled setting, where we generate a fully-synthetic dataset according to the click-model (Eq. 1). The synthetic data generation process samples observance and relevance probabilities from beta distributions: $\rho_{u,i} \sim Beta(1, 50)$ and $\gamma_{u,i} \sim Beta(0.5, 0.5)$; additionally, noise variables are sampled from a uniform distribution with parameter β: $\eta_{u,i} \sim Unif(0, \beta)$; and clicks are Bernoulli samples from the resulting click probabilities:

$$p(c_{u,i} = 1) = \min\left(\rho_{u,i} \cdot \gamma_{u,i} + \eta_{u,i},\ 1\right). \tag{18}$$

We generate the dataset with total number of users $|U| = 1000$, and total number of items $|I| = 100$. We randomly split the relevance R matrix into train and test sets, and generate clicks only from the train part of the split. To evaluate the robustness of the methods with respect to click noise, we vary the β parameter. For the sake of brevity, we only compare with MF-Dual as it is found to be the best performing baseline in the previous two settings.

6 RESULTS AND DISCUSSION

6.1 Semi-synthetic experimental results

The results on the unbiased relevance prediction task using the MovieLens-1M dataset are presented in Table 1. We evaluate the performance across different values of α, which controls the simulated selection bias, i.e., under different sampling test distributions. The VAE-IPS method consistently outperforms all methods by a significant margin across all metrics. The results hold for different settings of α, indicating the robustness of VAE-IPS across different degrees of selection bias.

Interestingly, Rel-MF and MF-DR perform worse than the vanilla MF model across all settings of α. We speculate that this is due to the biased negative loss in the Rel-MF formulation, as noted in previous work [20], and given that the MF-DR model is primarily aimed for explicit feedback data [44], it fails to perform in a binary feedback setting. VAE outperforms Rel-MF and MF-DR, possibly due to it

Table 1: Performance of different methods on the unbiased relevance prediction task on the MovieLens dataset. Significant improvements over MF-Dual are marked with † and over VAE with * ($p < 0.01$).

Method	$\alpha = 0.5$	$\alpha = 1.0$	$\alpha = 1.5$	$\alpha = 2.0$
	MAP@5			
MF	0.0336	0.0342	0.0343	0.0337
Rel-MF	0.0209	0.0212	0.0224	0.0239
MF-Dual	0.0489	0.0501	0.0491	0.0494
MF-DR	0.0236	0.0235	0.0239	0.0236
VAE	0.0399	0.0396	0.0399	0.0401
VAE-IPS	**0.0918**$^{\dagger *}$	**0.0914**$^{\dagger *}$	**0.0917**$^{\dagger *}$	**0.0909**$^{\dagger *}$
	NDCG@5			
MF	0.0198	0.0202	0.0203	0.0200
Rel-MF	0.0123	0.0125	0.0134	0.0141
MF-Dual	0.0289	0.0296	0.0292	0.0290
MF-DR	0.0138	0.0138	0.0140	0.0138
VAE	0.0236	0.0237	0.0236	0.0240
VAE-IPS	**0.0542**$^{\dagger *}$	**0.0536**$^{\dagger *}$	**0.0539**$^{\dagger *}$	**0.0535**$^{\dagger *}$

being a generative model, capable of capturing more complex patterns in the dataset. Nonetheless, VAE is still clearly outperformed by VAE-IPS by a large margin, as are the other baselines. Therefore, we conclude that, by combining the expressiveness of a generative model with a debiasing technique, VAE-IPS results in the highest performance across all tested settings.

6.2 Real-world experimental results

For the experimental setup on the Yahoo! R3 dataset, the results are presented in Table 2. Similar to the results on MovieLens-1M, VAE-IPS outperforms all other methods by a significant margin.

It is interesting to note that vanilla MF outperforms Rel-MF across all metrics in this dataset, even though the test-set is from a uniform random policy. We speculate that this is due to the estimation error in the propensity calculation [44]. Similar findings have been reported in previous work [33], where MF with IPS failed to outperform a vanilla MF method. Consistent with the results on the MovieLens-1M, MF-Dual outperforms all MF-based baselines and VAE without IPS. VAE-IPS clearly provides significantly higher performance than all other tested methods. Therefore, these results show us that, in addition to the semi-synthetic setting, the performance advantages of VAE-IPS are clearly observable on real-world data as well.

6.3 Fully-synthetic experimental results

To evaluate the robustness of VAE-IPS with respect to noise, we look at our experimental results on the fully-synthetic dataset in Figure 3. We vary the degree of noise, $\beta \in \{0.1, 0.3, 0.5, 0.7\}$, where higher values of β indicate higher levels of click noise (see Eq. 18); and for comparison, we also consider the performance of the MF-Dual, the highest performing baseline method in the previous settings.

Surprisingly, the performance of both MF-Dual and VAE-IPS is consistent across different noise values, with only small differences in performance between lowest ($\beta = 0.1$) and highest noise level ($\beta = 0.7$). This observation strongly indicates that they are both very robust to the noise in recorded clicks. Similar to what we have seen in our other experimental results, VAE-IPS has a higher performance across all metrics and noise settings. We thus conclude that VAE-IPS has a high robustness to noise and that it appears to outperform MF-Dual regardless of the level of click noise.

Table 2: Performance of different methods on a real-world dataset (Yahoo! R3). Significant improvements over MF-Dual are marked with † and over MF with * ($p < 0.01$).

Metric	MF	Rel-MF	MF-Dual	MF-DR	VAE	VAE-IPS
MAP@5	0.0212	0.0192	0.0212	0.0086	0.0198	**0.0234**$^{\dagger *}$
NDCG@5	0.0167	0.0151	0.0168	0.0068	0.0154	**0.0178**$^{\dagger *}$

Table 3: Performance of the linear (MF-Dual) and our VAE-IPS method on the fully-synthetic dataset, under varying degrees of noise: $\beta \in \{0.1, 0.3, 0.5, 0.7\}$.

Method	Metric	$\beta = 0.1$	$\beta = 0.3$	$\beta = 0.5$	$\beta = 0.7$
MF-Dual	MAP@5	0.2238	0.2261	0.2255	0.2541
	NDCG@5	0.3488	0.3517	0.3531	0.3451
VAE-IPS	MAP@5	**0.2316**	**0.2324**	**0.2348**	**0.2584**
	NDCG@5	**0.3615**	**0.3615**	**0.3654**	**0.3501**

7 CONCLUSION

In this paper we investigated whether state-of-the-art VAE recommendation models could be combined with debiasing techniques in the implicit feedback setting.

First, we studied the effect of bias on a naive VAE training objective based on clicks. Our analysis proved that directly optimizing this objective leads to a biased recommendation system that unfairly favors items that were overrepresented during data logging.

Second, we proposed VAE-IPS, a novel IPS correction for the VAE loss, that allows for the combination of VAE recommendation models with the IPS debaising method, and which is provenly unbiased w.r.t. selection bias in clicks. We evaluated VAE-IPS on two public datasets across various metrics and observed that it outperforms all our baselines across all metrics by significant margins. We believe our contribution of VAE-IPS is important to the recommendation systems field, because it combines the expressiveness of state-of-the-art VAE-based recommender models with IPS debiasing, and could lead to better performing recommender systems that are less affected by selection bias in interaction data.

Future work could consider robust methods for propensity estimation for implicit feedback in recommendation: IPS-based bias mitigation methods can be even more effective with more accurate propensity scores. IPS-based methods are known to suffer from the problem of high variance [9, 26], which can result in an *unsafe* policy and potentially lead to a negative user experience when deployed online. To avoid this, future work could consider adding safety regularization to the IPS objective [10], providing theoretical guarantees for *safe* deployment. Alternatively, future research can explore the application of other bias mitigation methods, such as the doubly-robust method [26, 44], to VAE recommendation models.

ACKNOWLEDGEMENTS

This research was supported by Huawei Finland and by the Hybrid Intelligence Center, a 10-year program funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research, https://hybrid-intelligence-centre.nl. This work used the Dutch national e-infrastructure with the support of the SURF Cooperative using grant no. EINF-4963. All content represents the opinion of the authors, which is not necessarily shared or endorsed by their respective employers and/or sponsors.

REFERENCES

[1] Aman Agarwal, Xuanhui Wang, Cheng Li, Michael Bendersky, and Marc Najork. 2019. Addressing Trust Bias for Unbiased Learning-to-rank. In *The World Wide Web Conference*. 4–14.

[2] Rocío Cañamares and Pablo Castells. 2018. Should I Follow the Crowd? A Probabilistic Analysis of the Effectiveness of Popularity in Recommender Systems. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. 415–424.

[3] Laurent Charlin, Rajesh Ranganath, James McInerney, and David M. Blei. 2015. Dynamic Poisson Factorization. In *Proceedings of the 9th ACM Conference on Recommender Systems*. 155–162.

[4] Aleksandr Chuklin, Ilya Markov, and Maarten de Rijke. 2015. *Click Models for Web Search*. Morgan & Claypool Publishers.

[5] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. 2008. An Experimental Comparison of Click Position-bias Models. In *Proceedings of the 2008 international conference on web search and data mining*. 87–94.

[6] Maurizio Ferrari Dacrema, Paolo Cremonesi, and Dietmar Jannach. 2019. Are We really Making Much Progress? A Worrying Analysis of Recent Neural Recommendation Approaches. In *Proceedings of the 13th ACM Conference on Recommender Systems*. 101–109.

[7] Marco Ferrante, Nicola Ferro, and Norbert Fuhr. 2021. Towards Meaningful Statements in IR Evaluation: Mapping Evaluation Measures to Interval Scales. *IEEE Access* 9 (2021), 136182–136216.

[8] Prem Gopalan, Jake M. Hofman, and David M. Blei. 2015. Scalable Recommendation with Hierarchical Poisson Factorization. In *Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence*. 326–335.

[9] Shashank Gupta, Philipp Hager, Jin Huang, Ali Vardasbi, and Harrie Oosterhuis. 2023. Recent Advances in the Foundations and Applications of Unbiased Learning to Rank. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

[10] Shashank Gupta, Harrie Oosterhuis, and Maarten de Rijke. 2023. Safe Deployment for Counterfactual Learning to Rank with Exposure-Based Risk Minimization. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

[11] F. Maxwell Harper and Joseph A. Konstan. 2015. The Movielens datasets: History and Context. *ACM Transactions on Interactive Intelligent Systems* 5, 4 (2015), 1–19.

[12] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. 2017. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web*. 173–182.

[13] Xiangnan He, Hanwang Zhang, Min-Yen Kan, and Tat-Seng Chua. 2016. Fast Matrix Factorization for Online Recommendation with Implicit Feedback. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. 549–558.

[14] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative Filtering for Implicit Feedback Datasets. In *Eighth IEEE International Conference on Data Mining*. IEEE, 263–272.

[15] Jin Huang, Harrie Oosterhuis, Maarten de Rijke, and Herke van Hoof. 2020. Keeping Dataset Biases Out of the Simulation: A Debiased Simulator for Reinforcement Learning Based Recommender Systems. In *Fourteenth ACM Conference on Recommender Systems*. 190–199.

[16] Ilija Ilievski and Sujoy Roy. 2013. Personalized News Recommendation Based on Implicit Feedback. In *Proceedings of the 2013 International News Recommender Systems Workshop and Challenge*. 10–15.

[17] Dietmar Jannach, Lukas Lerche, and Markus Zanker. 2018. Recommending Based on Implicit Feedback. In *Social Information Access*. Springer, 510–569.

[18] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. 2017. Unbiased Learning-to-rank with Biased Feedback. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. 781–789.

[19] Diederik P. Kingma and Max Welling. 2014. Auto-encoding Variational Bayes. In *International Conference on Learning Representations*.

[20] Jae-woong Lee, Seongmin Park, and Jongwuk Lee. 2021. Dual Unbiased Recommender Learning for Implicit Feedback. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1647–1651.

[21] Dawen Liang, Rahul G. Krishnan, Matthew D. Hoffman, and Tony Jebara. 2018. Variational Autoencoders for Collaborative Filtering. In *Proceedings of the 2018 world wide web conference*. 689–698.

[22] Roderick J.A. Little and Donald B. Rubin. 2002. *Statistical Analysis with Missing Data*. John Wiley & Sons.

[23] Benjamin M. Marlin and Richard S. Zemel. 2009. Collaborative Prediction and Ranking with Non-random Missing Data. In *Proceedings of the third ACM conference on Recommender systems*. 5–12.

[24] Benjamin M Marlin, Richard S Zemel, Sam Roweis, and Malcolm Slaney. 2007. Collaborative filtering and the missing at random assumption. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence*. 267–275.

[25] Harrie Oosterhuis. 2020. *Learning from User Interactions with Rankings: A Unification of the Field*. Ph. D. Dissertation. Informatics Institute, University of Amsterdam.

[26] Harrie Oosterhuis. 2023. Doubly Robust Estimation for Correcting Position Bias in Click Feedback for Unbiased Learning to Rank. *ACM Transactions on Information Systems* 41, 3 (2023), 1–33.

[27] Harrie Oosterhuis and Maarten de Rijke. 2021. Unifying Online and Counterfactual Learning to Rank: A Novel Counterfactual Estimator that Effectively Utilizes Online Interventions. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. 463–471.

[28] Harrie Oosterhuis, Rolf Jagerman, and Maarten de Rijke. 2020. Unbiased Learning to Rank: Counterfactual and Online Approaches. In *Companion Proceedings of the Web Conference 2020*. 299–300.

[29] Bruno Pradel, Nicolas Usunier, and Patrick Gallinari. 2012. Ranking with Non-random Missing Ratings: Influence of Popularity and Positivity on Evaluation Metrics. In *Proceedings of the sixth ACM conference on Recommender systems*. 147–154.

[30] Sanjay Purushotham, Yan Liu, and C-C Jay Kuo. 2012. Collaborative topic regression with social matrix factorization for recommendation systems. In *Proceedings of the 29th International Conference on Machine Learning*. 691–698.

[31] Yuta Saito. 2020. 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*. 309–318.

[32] Yuta Saito. 2020. Unbiased Pairwise Learning from Biased Implicit Feedback. In *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*. 5–12.

[33] Yuta Saito and Masahiro Nomura. 2022. Towards Resolving Propensity Contradiction in Offline Recommender Learning. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*. 2211–2217.

[34] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased Recommender Learning from Missing-not-at-random Implicit Feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 501–509.

[35] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *International Conference on Machine Learning*. PMLR, 1670–1679.

[36] Ilya Shenbin, Anton Alekseev, Elena Tutubalina, Valentin Malykh, and Sergey I Nikolenko. 2020. RecVAE: A New Variational Autoencoder for Top-n Recommendations with Implicit Feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining*. 528–536.

[37] Harald Steck. 2011. Item Popularity and Recommendation Accuracy. In *Proceedings of the fifth ACM conference on Recommender systems*. 125–132.

[38] Alex Strehl, John Langford, Lihong Li, and Sham M. Kakade. 2010. Learning from Logged Implicit Exploration Data. In *Advances in Neural Information Processing Systems*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta (Eds.), Vol. 23. Curran Associates, Inc.

[39] Xiaoyuan Su and Taghi M. Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques. *Advances in artificial intelligence* 2009 (2009).

[40] Adith Swaminathan and Thorsten Joachims. 2015. Batch Learning from Logged Bandit Feedback through Counterfactual Risk Minimization. *The Journal of Machine Learning Research* 16, 1 (2015), 1731–1755.

[41] Quoc-Tuan Truong, Aghiles Salah, and Hady W. Lauw. 2021. Bilateral Variational Autoencoder for Collaborative Filtering. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. 292–300.

[42] Ali Vardasbi, Harrie Oosterhuis, and Maarten de Rijke. 2020. When Inverse Propensity Scoring does not Work: Affine Corrections for Unbiased Learning to Rank. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 1475–1484.

[43] Xuanhui Wang, Nadav Golbandi, Michael Bendersky, Donald Metzler, and Marc Najork. 2018. Position Bias Estimation for Unbiased Learning to Rank in Personal Search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. 610–618.

[44] Xiaojie Wang, Rui Zhang, Yu Sun, and Jianzhong Qi. 2019. Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random. In *International*

Conference on Machine Learning. PMLR, 6638–6647.

[45] Sheng Zhang, Weihong Wang, James Ford, and Fillia Makedon. 2006. Learning from Incomplete Ratings Using Non-negative Matrix Factorization. In *Proceedings of the 2006 SIAM international conference on data mining*. SIAM, 549–553.

[46] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2019. Deep Learning Based Recommender System: A Survey and New Perspectives. *ACM Computing Surveys (CSUR)* 52, 1 (2019), 1–38.