VAE-IPS: A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback

SHASHANK GUPTA, University of Amsterdam, The Netherlands

HARRIE OOSTERHUIS, Radboud Universiteit, The Netherlands

MAARTEN DE RIJKE, University of Amsterdam, The Netherlands

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. In this work, we address this gap by introducing an inverse propensity scoring (IPS) based unbiased training method for VAEs from implicit feedback data, VAE-IPS, which 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.

CCS Concepts: • **Information systems** → *Recommender systems.*

Additional Key Words and Phrases: Variational autoencoder; Implicit feedback

ACM Reference Format:

Shashank Gupta, Harrie Oosterhuis, and Maarten de Rijke. 2022. VAE-IPS: A Deep Generative Recommendation Method for Unbiased Learning from Implicit Feedback. In *CONSEQUENCES+REVEAL Workshop at RecSys '22, September 22–23, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 6 pages.

1 INTRODUCTION

Recommender systems rely on user feedback signals to infer user preferences [7, 24]. It is well-known that 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 [9, 10, 15, 17]. To mitigate the negative effects of selection bias, existing work has proposed the usage of inverse propensity scoring (IPS), a counterfactual estimation technique [21]. Recently, the IPS approach has been extended to optimize matrix factorization (MF) models from implicit feedback [20]. While MF methods [6, 7], and the more recent neural MF-based methods [5, 29], have a long tradition in the recommendation field, state-of-the-art methods for learning from implicit feedback data use variational autoencoders (VAEs) instead [13, 22, 25]. Despite the importance of bias mitigation on the one hand, and the strong performance of VAEs for recommendation from implicit feedback on the other hand, existing work has overlooked the issue of bias in a generative setting [12, 20, 21]. We call attention to this problem and address this gap in the literature by introducing VAE-IPS, an IPS debiasing method for VAE optimization from implicit feedback that optimizes the ideal generative objective in expectation. Specifically, our main contribution is to model generative user-item relevance in an unbiased fashion, building on the BiVAE framework [25].

2 METHOD: VAE-IPS ESTIMATOR

In this section, we introduce the user behavior model that we use, the ideal generative objective for relevance, the naive click-based objective, before defining our IPS-corrected unbiased generative VAE-IPS estimator.

Click model. We assume a simple click model, where the click probability ($P(c_{u,i} = 1)$) is a product of the probability of observance ($o_{u,i}$) and that of relevance ($r_{u,i}$) [2, 16, 19, 20, 26]: $P(c_{u,i} = 1) = P(o_{u,i} = 1) \cdot P(r_{u,i} = 1) = \rho_{u,i} \cdot \gamma_{u,i}$, where $\rho_{u,i}$ and $\gamma_{u,i}$ are the observation and relevance probability, respectively.

Variational autoencoder for click feedback. The *ideal* relevance generative objective can be described as follows:

$$L^{rel}_{u,i} \geq \underbrace{\mathbb{E}_{q_\phi(z_{u,i})} \log\left[p_\theta(r_{u,i}|z_{u,i})\right]}_{\textit{Conditional likelihood}} - \underbrace{D\left(q_\phi(z_{u,i}) \| p(z_{u,i})\right)}_{\textit{KLD-regularizer}} = L^{ideal}_{u,i}, \tag{1}$$

where we define a lower-bound of the log-likelihood, $L^{ideal}_{u,i}$, also known as the evidence lower bound objective (ELBO) in the autoencoder literature [11, 25]. Here, $q_\phi(z_{u,i})$ is the posterior distribution over the user-item latent variable $z_{u,i}$. This is the *ideal* distribution, since the ELBO is the quantity that is optimized in VAEs instead of the log-likelihood ($L^{rel}_{u,i}$) [11].

Given that relevance is a binary random variable for implicit feedback data ($r_{ui} \in [0, 1]$), plugging it into Eq. 1, we obtain:

$$L^{ideal}_{u,i} = \mathbb{E}_{q_\phi(z_{u,i})}\left[r_{ui} \log(\pi_\theta(z_{u,i})) + \ (1 - r_{u,i}) \log(1 - \pi_\theta(z_{u,i}))\right] - D\left(q_\phi(z_{u,i}) \| p(z_{u,i})\right), \tag{2}$$

where $\pi_\theta(z_{u,i})$ is the relevance score for the pair (u, i). For the sake of brevity, we refer the reader to [25] for a more detailed treatment of the ELBO. BiVAE makes use of clicks, instead of relevance, which results in the following loss function:

$$L^{click}_{u,i} = \mathbb{E}_{q_\phi}\left[c_{ui} \log(\pi_\theta(z_{u,i})) + (1 - c_{u,i}) \log(1 - \pi_\theta(z_{u,i}))\right] - D\left(q_\phi(z_{u,i}) \| p(z_{u,i})\right). \tag{3}$$

This is a biased estimator, where the bias w.r.t. $o_{u,i}$, can be defined as $\mathbb{E}_o[L^{click}_{u,i}]$:

$$\mathbb{E}_o[L^{click}_{u,i}] - L^{ideal}_{u,i} = \mathbb{E}_{q_\phi}\left[(\rho_{u,i} - 1) r_{u,i} \log\left(\frac{\pi_\theta(z_{u,i})}{1 - \pi_\theta(z_{u,i})}\right)\right]. \tag{4}$$

From Eq. 4 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. Next, we will introduce our proposed estimator VAE-IPS, which corrects for this bias using IPS.

2.1 Proposed estimator

We propose an unbiased generative estimator for VAE, in a similar vain as existing IPS corrections for existing biases [1, 8, 27]. Our proposed estimator VAE-IPS, an unbiased estimate of the true generative objective, is defined as follows:

$$L^{ips}_{u,i} = \mathbb{E}_{q_\phi}\left[\frac{c_{u,i}}{\rho_{u,i}} \log(\pi_\theta(z_{u,i})) + \left(1 - \frac{c_{u,i}}{\rho_{u,i}}\right) \log(1 - (\pi_\theta(z_{u,i})))\right] - D\left(q_\phi(z_{u,i}) \| p(z_{u,i})\right). \tag{5}$$

This estimator is an unbiased estimate of the true relevance based objective (Eq. 2). To prove this, we derive the expected value of the estimator with respect to the observation variable:

$$\mathbb{E}_o\left[L^{ips}_{u,i}\right] + D\left(q_\phi(z_{u,i}) \| p(z_{u,i})\right) = \mathbb{E}_{o,q_\phi}\left[\frac{c_{ui}}{\rho_{u,i}} \log(\pi_\theta(z_{u,i})) + \left(1 - \frac{c_{ui}}{\rho_{u,i}}\right) \log(1 - \pi_\theta(z_{u,i}))\right] \tag{6}$$

$$= \mathbb{E}_{q_\phi}\left[\mathbb{E}_o\left[\frac{c_{ui}}{\rho_{u,i}}\right]\log(\pi_\theta(z_{u,i})) + \left(1 - \mathbb{E}_o\left[\frac{c_{ui}}{\rho_{u,i}}\right]\right)\log(1-\pi_\theta(z_{u,i}))\right] \tag{7}$$

$$= \mathbb{E}_{q_\phi}\left[r_{u,i}\log(\pi_\theta(z_{u,i})) + (1-r_{u,i})\log(1-\pi_\theta(z_{u,i}))\right]. \tag{8}$$

Thus, in expectation it is equal to the ideal relevance-based objective from Eq. 2. This proves that the introduced VAE-IPS estimator is unbiased.

3 EXPERIMENTAL RESULTS

We assess the performance of VAE-IPS using the unbiased relevance prediction task, with a real-world and a semi-synthetic setup. We use the following baseline methods: (i) **Binary Matrix Factorization**: We use the matrix factorization model for implicit feedback dataset from [7], 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 [20]. (iii) **MF-DR**: A doubly-robust variant of the IPS matrix factorization model, which uses a control variate to reduce the variance of the IPS method [28]. (iv) **MF-DU**: 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 [12]. (v) **VAE**: We use the BiVAE framework developed to model dyadic data [25], which is more suitable for pointwise predictions. This VAE baseline is optimized with the proposed alternate coordinate descent style optimization method, where the posteriors for user and items are optimized alternately. (vi) **VAE-IPS**: This is our proposed method, the BiVAE model optimized with the unbiased VAE-IPS objective (Eq. 5). 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. For variance reduction we apply propensity clipping [20, 23], and for propensity estimation we use the item's relative click frequency in the training dataset and make the assumption that propensity scores are uniform across all users [20].

3.1 Experimental setup

Real-world dataset experimental setup. We 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 [14], 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.

Semi-synthetic experimental setup. We als evaluate VAE-IPS using the MovieLens-1M dataset [4]. 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 [18]. 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 the 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.

Further details. To get a validation set, we split the training set in both of the 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 [21]. We use NDCG@k and MAP@k as evaluation metrics with varying cut-off lengths ($k = 1, 3, 5$). For calculating the normalizing factor of the DCG@k, we follow the advice from [3]. Our evaluation metrics follow the definitions of earlier work on Rel-MF [20].

Table 1. Performance of different methods on real-world the dataset (Yahoo! R3). Significant improvements over the baseline (MF-DU) are marked with † ($p < 0.01$). Reported numbers are all percentages (%).

	MAP			NDCG		
Method	@1	@3	@5	@1	@3	@5
MF	0.502	0.977	1.235	0.502	0.830	0.997
MF-DR	0.431	0.752	0.925	0.431	0.615	0.720
Rel-MF	0.678	1.333	1.634	0.678	1.136	1.312
MF-DU	0.890	1.638	2.013	0.890	1.367	1.588
VAE	0.829	1.534	1.887	0.829	1.282	1.493
VAE-IPS	**1.101**†	**1.949**†	**2.356**†	**1.101**†	**1.603**†	**1.828**†

4 RESULTS AND DISCUSSION

For the real-world experimental setting, using the Yahoo! R3 dataset, the results are presented in Table 1. VAE-IPS outperforms all other methods by a significant margin. It is interesting to note that Rel-MF is outperforming vanilla MF across all metrics in this dataset. We speculate that this is due to the test set coming from an unrealistic truly uniform random logging policy, where the assumptions of Rel-MF hold. Consistent with the results on the MovieLens-1M on which we report below, MF-DU outperforms all MF-based baselines and VAE without IPS. VAE-IPS clearly provides significantly higher performance than all other tested methods, across all metrics.

Due to space constraints, the results for the unbiased relevance prediction task on the semi-synthetic experimental setting using the MovieLens-1M dataset are presented in Appendix A (Table 2). Again, 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 w.r.t. different degrees of selection bias. A higher α value corresponds to a higher selection bias; and for this setting, most of the baseline methods' performances drop considerably, whereas the performance of VAE-IPS is only moderately reduced. Interestingly, the performance of the baseline MF-DU is consistent across different settings of α. We speculate that this is because of the unbiased negative click loss in the MF-DU model, as opposed to a biased negative click loss in Rel-MF [12]. Given that the majority of clicks in a real-world setting are negative, with an unbiased negative loss, MF-DU performs better than the other baselines. Interestingly, Rel-MF and MF-DR perform worse than the vanilla MF model across all settings of α. We speculate that this is because of the biased negative loss in the Rel-MF formulation and the MF-DR model formulation primarily aimed for explicit feedback data. VAE outperforms Rel-MF, and MF-DR, possibly due to it being a generative model, capable of capturing more complex patterns in the dataset. Nonetheless, it is still clearly outperformed by VAE-IPS.

5 CONCLUSION

In this paper, we addressed an important gap in the literature on unbiased recommendation: bias in a generative VAE model for implicit feedback. We investigated the application of IPS techniques for debiasing the state-of-the-art VAE recommendation models in the implicit feedback setting, viz. VAE-IPS, a novel IPS correction for the VAE loss. Our proposed method allows for the combination of VAE recommendation models with the IPS debiasing method, and is provably unbiased w.r.t. selection bias in clicks. We evaluated VAE-IPS on two public datasets across various metrics and observed that it outperforms every baseline across all metrics by a significant margin. Future work could consider propensity estimation for implicit feedback in recommendation: a limitation of existing IPS methods is that they require accurate propensity scores, and thus errors in propensity estimation can propagate to later debiasing steps.

ACKNOWLEDGMENTS

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

[3] 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.

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

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

[10] 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.

[11] Diederik P Kingma and Max Welling. 2013. Auto-encoding Variational Bayes. *arXiv preprint arXiv:1312.6114* (2013).

[12] 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.

[13] 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.

[14] 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.

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

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] Xiaoyuan Su and Taghi M Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence* 2009 (2009).

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

A SEMI-SYNTHETIC EXPERIMENTAL RESULTS

The results from our semi-synthetic experimental setup, with different experimental settings (α values) are reported in this section.

Table 2. Performance of different methods on the unbiased relevance prediction task on the MovieLens dataset. Significant improvements over MF-DU are marked with † ($p < 0.01$), and over VAE are marked with *. Reported numbers are all percentages (%).

Exp. setting	Method	MAP@1	MAP@3	MAP@5	NDCG@1	NDCG@3	NDCG@5
$\alpha = 0.5$	MF	3.036	5.441	6.667	3.036	3.831	3.899
	MF-DR	0.888	1.588	2.014	0.888	1.117	1.191
	Rel-MF	2.902	5.160	6.377	2.902	3.613	3.725
	MF-DU	2.492	4.607	5.783	2.492	3.278	3.436
	VAE	4.183	7.493	9.234	4.183	5.273	5.408
	VAE-IPS	**5.434**†*	**9.903**†*	**12.202**†*	**5.434**†*	**7.015**†*	**7.181**†*
$\alpha = 1.0$	MF	1.097	1.960	2.470	1.097	1.376	1.456
	MF-DR	0.832	1.510	1.921	0.832	1.067	1.140
	Rel-MF	1.157	1.992	2.381	1.157	1.383	1.366
	MF-DU	2.562	4.686	5.872	2.562	3.321	3.476
	VAE	1.687	3.137	3.944	1.687	2.237	2.347
	VAE-IPS	**3.885**†*	**7.090**†*	**8.717**†*	**3.885**†*	**5.023**†*	**5.128**†*
$\alpha = 1.5$	MF	0.326	0.652	0.866	0.326	0.477	0.532
	MF-DR	0.865	1.580	2.014	0.865	1.119	1.198
	Rel-MF	0.255	0.551	0.731	0.255	0.412	0.457
	MF-DU	2.482	4.615	5.773	2.482	3.292	3.432
	VAE	0.539	1.191	1.582	0.539	0.896	0.994
	VAE-IPS	**3.344**†*	**6.123**†*	**7.516**†*	**3.344**†*	**4.345**†*	**4.424**†*
$\alpha = 2$	MF	0.189	0.299	0.379	0.189	0.200	0.216
	MF-DR	0.865	1.585	1.986	0.865	1.126	1.177
	Rel-MF	0.055	0.169	0.245	0.055	0.139	0.165
	MF-DU	2.478	4.605	5.761	2.478	3.282	3.423
	VAE	0.292	0.635	0.818	0.292	0.476	0.509
	VAE-IPS	**3.151**†*	**5.631**†*	**6.918**†*	**3.151**†*	**3.959**†*	**4.044**†*