Modeling Variational Anchoring Effect for Recommender Systems

Yudi Xiao
School of Economics and Management
Institute for Advanced Intelligence
Dalian University of Technology
Dalian, China
xiaoyd0220@mail.dlut.edu.cn

Yingyi Zhang
School of Economics and Management
Institute for Advanced Intelligence
Dalian University of Technology
Dalian, China
yingyizhang@mail.dlut.edu.cn

Xianneng Li*
School of Economics and Management
Institute for Advanced Intelligence
Dalian University of Technology
Dalian, China
xianneng@dlut.edu.cn

***Abstract*—Users generally have a tendency to rely on numerical information of recommendations presented on the web page when judging the recommended items, which refers to a classic psychological concept, anchoring effect. Learning users' psychology from explicit behaviors has been widely applied in RS and performs well on capturing user preferences and guiding the prediction tasks of recommendations. Recent studies have empirically proven that the anchoring effect can mislead users to click/purchase items that are not liked in principle, which will bring bias and noise to behavior data. However, vast majority of existing recommendation algorithms trained on behavior data ignore the anchoring bias, which results in suboptimal recommendations. In this paper, we propose a novel method named Variational Anchoring Effect Encoder (VAEE) to model the anchoring effect and mitigate the anchoring bias for recommender systems. The proposed method mainly includes two steps: 1) User Anchoring Effect Module which aims to reconstruct the unanchored user preferences with a Variational Autoencoder (VAE)-based deep structure, and 2) User Anchoring Debias Module that generate the recommendation results with the reconstructed unbiased user representations. Extensive experiments on real-world datasets are conducted to demonstrate that reducing anchoring effect can bring particular improvement in AUC and the proposed VAEE is attachable to most existing recommendation models. We also compare the recommendation quality when using different anchoring feature subsets, which indicates that the learned representation of anchoring effect is authentic and truly effective to restore users' true preferences.**

Keywords—anchoring bias, anchoring effect, recommender systems, variational autoencoder

I. INTRODUCTION

Recommender system (RS) is one of the most prevalent tools for online retailers to gain popularity and promote sales [1] as it provides users with products closest to their preferences. With the help of RS, consumers can easily find ideal items and even possible to discover their hidden interests. Most of the recommendation algorithms today use consumer behavior (e.g. click or purchase) from log data as the main input to make predictions [1-3], since consumer behaviors are widely agreed to reflect their psychology and preferences, providing sufficient information for recommendations.

However, research has found that recommendations can cause a psychological phenomenon called **anchoring effect**, which indicates that consumers are often unconsciously misled by the numerical information provided with recommendations [4]. For example, if the presented prices of most recommended items are around $500, the price of the item that the user purchased will become closer to $500. Hence, anchoring effect can cause a serious bias in consumers' decision-making process and reshape their behaviors (e.g. click or purchase) [5-6].

Users tend to temporally 'prefer' the recommended items, which means their decisions will depend more on the information of these items. This gives birth to more unintended behavior records when users interact with items that are actually not their cup of tea but only recommended.

When making recommendation predictions, these biased behavior data will draw more attention to useless information and interfere with the learning process of true preferences. Moreover, preferences and recommendations will become more similar to each other. Each prediction that the recommender made only attempts to prove the correctness of itself, leaving aside the true requirement of users. In a word, consumer behavior can be severely biased and fail to represent consumers' true preferences [7-8], which can impair the effectiveness of RS. For better recommendations, the anchoring bias should be specifically examined and carefully removed under closer investigation.

Although numerous studies have concentrated on data debias in RS [9-11], the anchoring bias remains an inadequately-researched problem. Investigations on anchoring effect in RS concentrate mostly on revealing the existence of the effect while lacking a clear and explicit definition and quantification of the anchoring bias. These researches leave behind two tough issues in solving the anchoring problem:

1) *Anchoring Representation*: how to separate the anchoring effect from the users' true preferences. As mentioned above, user preferences can be misled by unexpected numerical features. Thus, what the RS needs is to restore user unbiased preferences, which requires generating anchoring effect first.

2) *Anchoring Mitigation*: how to relieve the anchoring bias in RS. Followed by the representation of anchoring effect, the recommendation algorithm can be trained by unanchored user preferences where anchoring effect can be mitigated.

Hence, in this paper, we propose a novel method named **V**ariational **A**nchoring **E**ffect **E**ncoder (VAEE) to model anchoring effect and mitigate the anchoring bias. It mainly includes two steps: **1) User Anchoring Effect Module** and **2)**

*Corresponding author: Xianneng Li. This research is supported by the National Natural Science Foundation of China (NSFC) under Grant 72071029 and 72231010.

DOI 10.1109/CAI59869.2024.00170

User Anchoring Debias Module. Firstly, **User Anchoring Debias Module** aims to reconstruct the unanchored user preferences with a Variational Autoencoder (VAE)-based deep structure. VAE has been examined as an outstanding generative structure for striping the latent distribution from features [12-14], and widely incorporated into collaborative filtering to capture latent information [15-17]. This enables VAEE to represent the latent distribution of anchoring effect if appropriately designed. In the second step, **User Anchoring Debias Module** generates the recommendation results with unbiased user representations. This provides a practicable path to construct an anchoring representation to improve the recommendation performance. After modeling anchoring effect, the unanchored user preferences can be acquired by mitigating anchoring effect to train on traditional RS like DSSM [18], NMF [3], and DeepFM [2] to achieve improved performance.

The summary of our contributions goes as follows:

- To the best of our knowledge, this is the first solution to mitigate anchoring bias in RS that has been addressed before. We design a novel framework with a developed VAE module to reduce the negative influence of anchoring effect and reconstruct users' unanchored preferences.
- We extend the application of VAE in RS for debias. We provide a VAE-based module that can be easily used in other RS for mitigating anchoring effect.
- We conduct offline experiments in several large e-commerce datasets, proving that our method can significantly improve traditional RS.

II. Model

The structure of the proposed VAEE is shown in Fig. 1 and Fig. 2, which includes two main components. Fig. 1 represents the User Anchoring Effect Module, which demonstrates how to learn the anchoring representation via VAE. Fig. 2 is the User Anchoring Debias Module that illustrates how to eliminate the anchoring bias for recommendations.

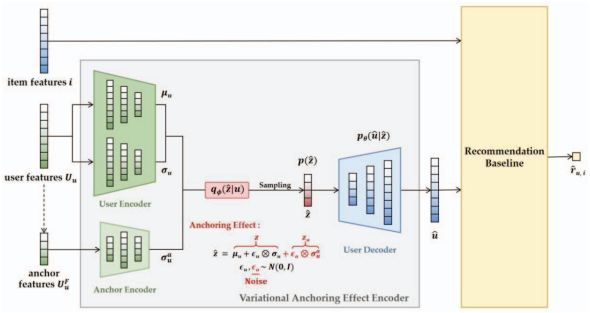

Fig. 1. VAEE with User Anchoring Effect Module

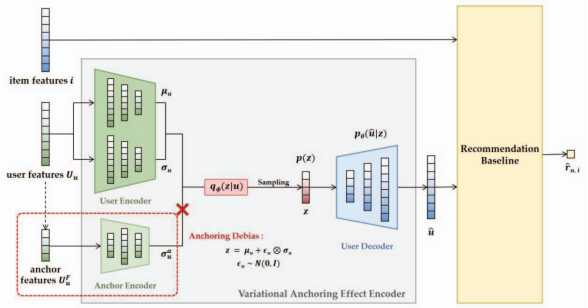

Fig. 2. VAEE with User Anchoring Debias Module

Basic Notations. We assume a set of users $\mathbf{U} = \{U_1, U_2, \dots\}$ where U_u represents meta features of user u. And a set of items $\mathbf{I} = \{i_1, i_2, \dots\}$ that users in U have interacted with. $r_{u,i} \in (0,1)$ denotes the label of each user-item interaction. Some user features related with anchoring effect are selected according to former research [5,19] to create a feature subset $\mathbf{F}$ (e.g. $\mathbf{F} = \{\text{id}, \text{gender}, \text{age}, \dots\}$), and the user anchoring features are represented as $\mathbf{A} = \{U_1^{\mathbf{F}}, U_2^{\mathbf{F}}, \dots\}$. Details about the selection of anchor-related features will be discussed in experiment section.

VAE Model. Variational AutoEncoder(VAE) is a probabilistic generative architecture with a prior and noise distribution, respectively [20], whose objective is to maximize the likelihood of input data via an encoder-decoder manner. The encoder uses a neural network to map the input into a low-dimensional latent space while the decoder oppositely maps the latent sample back to the input space. Between the encoder and decoder, a latent representation is randomly sampled using a reparameterization trick to allow back-propagation for model parameters. To optimize the model, the loss function combines reconstruction error and Kullback-Leibler divergence for better performance on target reconstruction from a random Gaussian posterior distribution.

A. User Feature Reconstruction

Since anchoring effect refers to a psychological phenomenon [4,7,21], it naturally lack a direct and explicit representation from user data. So it's necessary to clearly define a proper and reasonable representation to quantify the anchoring effect before discussing how to mitigate it. According to former researches, some of the user features are claimed to be significant factors to anchoring effect [19,22]. Hence, we manage to capture anchor-related information from user metadata by reconstructing user features with a VAE generative structure and generating an anchoring representation for each user.

For the target encoder of VAE, user meta features U_u serve as the input to learn the parameters of the encoding variational distribution, i.e., the mean value μ_u and the standard variance σ_u of a Gaussian distribution, as follows:

$$\mu_u = ReLU(W_\mu U_u + b_\mu), \tag{1}$$

$$\sigma_u = exp(ReLU(W_\sigma U_u + b_\sigma)), \tag{2}$$

where W_μ, W_σ and b_μ, b_σ are the weight matrices and bias vectors, and $ReLU(\cdot)$ is the activation function.

With the mean μ_u and the standard deviation σ_u, the latent representation z of user features can be sampled from the encoding variational distribution:

$$q_\phi(z|u) = N(\mu_u, \sigma_u^2 I), \tag{3}$$

However, random sampling can block the back-propagation of training gradients. To solve this, VAE applied a reparameterization trick [20] as follows to approximate the random sample z of the latent representation from the encoding distribution:

$$z = \mu_u + \epsilon_u \otimes \sigma_u, \epsilon_u \sim \mathrm{N}(0, I), \tag{4}$$

where $\otimes$ denotes the element-wise product of vectors.

For the decoder of VAE, the representation sample z goes through decoding neural layers with the reversed design of the encoding one, and generates a reconstructed representation $\hat{u}$ from the decoding distribution $p(z|\hat{u})$ for each user:

$$\hat{u} = ReLU(W_z z + b_z), \tag{5}$$

where W_z and b_z are the weight matrices and bias vectors.

User Anchoring Effect Module. Anchoring effect can cause user preferences bias towards the observed attributes (can be also denoted as explicit features) of recommended items. To construct a latent representation for anchoring effect, we denote the anchoring bias as z_a and assume $z_a \sim \mathrm{N}(0, (\sigma_u^a)^2 I)$, where an additional neural module is designed to learn the distribution from the anchoring features **F**:

$$\sigma_a^u = exp(ReLU(W_a U_u^{\mathbf{F}} + b_a)). \tag{6}$$

Since the latest research on the mechanism of anchoring effect [23-24] argues that its impact is also largely associated with environmental noise, we additionally sample a random variable ϵ_a from standard normal distribution $\mathrm{N}(0, I)$ to simulate the uncertain environmental conditions. Afterwards, we add the representation of anchoring bias into the latent distribution as follows so that we can learn it through the whole training process:

$$\hat{z} = z + z_a = \mu_u + \epsilon_u \otimes \sigma_u + \epsilon_a \otimes \sigma_u^a, \tag{7}$$

Combining the anchoring representation, the posterior distribution learned by the encoder as in (3) changes into:

$$q_\phi(\hat{z}|u) = N(\mu_u, \sigma_u^2 + (\sigma_a^u)^2 I), \tag{8}$$

and the input z of the reconstructed user representation $\hat{u}$ in (5) changes into $\hat{z}$.

By this means, VAEE separates the anchoring effect from user preferences and learns an explicit representation of the anchoring bias. This offers the opportunity to mitigate the negative bias so that VAEE can learn user preferences without the interference of anchoring effect.

User Anchoring Debias Module. To mitigate anchoring effect, the anchoring representation should be removed before predictions so our final recommendations can be less biased. However, we can only learn anchor representation during the prediction because anchoring effect derives from the personalized product recommendations for each user. This indicates the difficulty to learn proper parameters for anchoring representation during common training process.

So we design a new learning strategy to resolve this difficulty. The training stage can be seen in Fig. 1, whereas the inference stage can be illustrated in Fig. 2. During training, the entire model aims to learn the latent representation of anchoring effect and train ideal recommendations simultaneously. While for validations, we drop the learned anchoring representation in VAE to remove the influence of anchoring effect, which indicates that the final recommendations can be less biased. The total loss function of training and validation is adjusted according to this strategy, which will be introduced in the next subsection.

III. Optimization

According to original VAE assumptions, the learned posterior distribution is supposed to be similar to the assumed prior distribution $p(z) \sim N(0, I)$, which can be measured through KL divergence (L_{KL}). Besides, the target of user VAE is to maximize the similarity between true user representation and the reconstructed user representation (L_{rec}). So the loss function of VAE combines these two target functions:

$$loss_{VAEE} = L_{rec} + L_{KL}, \tag{9}$$

$$L_{rec}(u, \hat{u}) = E_{q_\phi(\hat{z}|u)}[log(p_\theta(\hat{u}|\hat{z}))], \tag{10}$$

$$L_{KL}(\mu_u, \sigma_u^2) = KL(q_\phi(\hat{z}|u) \| p(\hat{z})), \tag{11}$$

where $E_{\cdot}[\cdot]$ refers to expectation, $KL(\cdot)$ refers to KL divergence.

Combining the reconstructed user features and item features, the recommender baseline makes final predictions for user-preferred items, and the loss function $loss_{pred}$ is based on the widely-used log-likelihood function:

$$loss_{pred} = -\sum_{u \in U,\, i \in I} r_{u,i}\, log(\mathrm{P}(\hat{r}_{u,i}|\hat{u}, i)). \tag{12}$$

According to our training strategy (in the previous section), the total loss function is respectively adjusted for training and validation (testing) as follows to achieve the goal of representing and mitigating anchoring effect:

$$loss = loss_{pred} + loss_{VAEE} \quad (\text{for Training}), \tag{13}$$

$$loss = loss_{pred} \quad (\text{for Validation and Testing}). \tag{14}$$

The novel anchoring-processing modules can be regarded as a further feature engineering to remove the underlying anchoring factors within user metadata that can impact the recommendation quality. Afterwards, the user representation without the anchoring bias can be incorporated into a RS for recommendation, where most of the existing representatives can be embedded. In this paper, we select three baselines to demonstrate the applicability of the proposed VAEE, including the traditional NMF, and two deep learning methods, i.e., DSSM [18] and DeepFM [2], for experiments.

IV. EXPERIMENTS

In this section, two experiments are described in detail. The first experiment compares the recommendation performance of the proposed VAEE using three base models to demonstrate its effectiveness. The second one illustrates the existence of anchoring effect by comparing different combinations of the selected anchoring features. The common goal of both experiments is to provide convincing evidence that the proposed VAEE can effectively mitigate the anchoring effect in recommendations.

A. Experimental Setup

Datasets. To test the recommendation quality of baselines and the proposed VAEE, three open-source real-world datasets are selected for experiments. Table I shows statistics of the evaluation datasets.

TABLE I. STATSTICS OF THE DATASETS

Dataset	#Users	#Items	#Interactions
CIKM 2019 Ecomm AI	80,000	912,114	3,234,367
Amazon-Electronics	9,560	1,157,633	1,292,954
Movielens-1M	6,040	3,706	1,000,209

- CIKM 2019 Ecomm AI[1]: This dataset contains 3,234,367 pieces of purchase records of 912,114 items created by 80,000 users and other auxiliary information, including gender, age, and occupation for users, while category, shop, and brand for items.
- Amazon-Electronics[2] [25]: The dataset contains 1,292,954 ratings of 1,157,633 electronic products given by 9,560 users. Note that this dataset only has gender as user meta features, while product category, first year on sale, and brand for products.
- Movielens-1M[3] [26]: The dataset has 1,000,209 ratings of 3,706 movies given by 6,040 users. User meta features include gender, age, and occupation, while item meta features include movie title and genre.

Details about dataset processing and related settings are summarized as follows:

1) Duplicated interactions. Duplicate records of the same user-item pair are integrated into one piece, keeping the information of the latest record only.

2) Sampling. The proportion of positive and negative instances is 1:1. For behavior datasets, behavior records serve as positive instances, and negative instances are randomly sampled from unobserved interactions. For rating datasets (e.g., ratings between 1 and 5), ratings of 3 or higher are positive instances and ratings of 2 or lower become negative instances.

3) Recommendation task. For CIKM dataset, the task is to predict whether the user will purchase the target item. For Amazon and Movielens datasets, the task is to figure out user preferences for the target item (whether the user will give a high rating).

1 https://tianchi.aliyun.com/competition/entrance/231721/introduction
2 https://cseweb.ucsd.edu/ jmcauley/datasets.html
3 https://grouplens.org/datasets/movielens/

Baselines. DSSM [18], DeepFM [2], and NMF [3] are respectively used as base models of the proposed VAEE. We conduct comparing experiments to illustrate the improvement that our debiasing framework has brought for these classic recommendation models.

- DSSM [18] uses basic MLP to project user-item pairs(meta features) into a common latent space and measure their semantic similarity with a cosine method.
- NMF [2] combines Factorization Matrix with Multi-layer Perceptron, to learn non-linear latent representations for user/item features and model the two-way interactions of user-item pairs.
- DeepFM [3] is an integration of DNN and Factorization Machine, that DNN learns the non-linear and high-order latent information from features and FM captures the linear and low-order feature interactions

Each baseline is trained in three framework settings to further examine the effectiveness of each module in the proposed model. *BaseModel* refers to the baseline trained alone. Compared with *BaseModel*, *BaseModel+VAE* uses the VAE structure to reconstruct user representations without considering the existence of anchoring effect. Moreover, *BaseModel+VAEE* uses the VAEE structure and trains the model following the proposed new training method that mitigates anchoring effect for final recommendations. The results are shown in Table II.

Anchoring features. Former empirical research on consumer psychology has presented convincing evidence that gender can be a significant impact factor with anchoring effect [19,22]. So gender is selected to train the anchoring module. Moreover, as users are influenced by anchoring effect in different degrees, user-id is also concatenated into the input of anchoring representations to learn the peculiarity of each user. Note that since most open-source datasets lack various user features, our attempt to learn the anchoring representation is somewhat limited, remaining the further in-depth investigation beyond this study.

B. Experiments with Real-world Datasets

Parameter settings. The recommendation methods are implemented based on Pytorch. We tested the batch size of [1024, 2048, 4096], and the learning rate of [0.0005, 0.0001, 0.005]. The embed size for each feature is 32.

We test the optimal parameter settings of each base model for better recommendation performance on the three datasets. For DSSM, the network structure of the DNN layers for user and item semantic features is (256,128). For NMF, the network structure to learn the latent interactions of user-item pairs is (128,32,8). For DeepFM, the structure of deep hidden layers is (64,8). Other parameters are set according to the original model design in [2,3,18].

For fairness, the parameters of base models used in our VAEE remain the same as the ones which are trained alone. The encoder in VAE is structured as 256-128 and the decoder is structured in reverse. As for activation functions, Sigmoid function is used for output layers and ReLU for other layers. The optimizer is Adaptive Moment Estimation (Adam) [27] and each experiment has been repeated 5 times.

Performance evaluation. Main evaluation metric is AUC. It refers to the Area Under the Receiver Operating Characteristic Curve (ROC). The calculation follows:

$$AUC = \frac{SumRank}{M \times N}, \quad (15)$$

where *SumRank* refers to the sum of Indicator function output of correctly predicted pairs of positive samples (i.e. $\hat{r}_{u,i} = 1$) and negative samples, *M* is the number of positive samples and *N* is the number of negative samples.

Table II demonstrates the results of the offline experiments . For each baseline, *BaseModel+VAE* and *BaseModel+VAEE* outperform the *BaseModel* in user preferences predictions. The average improvement of VAEE is 2.772% (from 0.278% to 5.610%). This demonstrates that using the VAE structure to reconstruct user representations can be helpful for a better understanding of user preferences, showing the effectiveness of the framework design. Since no other works rebuilding user/item features with VAE have been presented, our achievements show a promising idea for more attempts to extend the application of VAE in recommendation algorithms.

In most cases, *BaseModel+VAEE* can reach better performance than *BaseModel+VAE*, which demonstrates that the proposed training method is effective in mitigating the negative influence of anchoring effect. Although the improvement from VAE to VAEE is around 0.1%, the adjunction of the VAEE module still assists original recommendation baselines in learning users' true preferences better. This confirms our assumption that there are some hidden factors within the user input that are related to anchoring bias, which will be detailedly discussed in the next section.

TABLE II. AUC OF OFFLINE COMPARING EXPERIMENTS ON TEST DATASETS

Models	CIKM 2019 Ecomm AI	Amazon-Electronics	Movielens-1M
DSSM	0.808	0.843	0.787
DSSM+VAE	0.823*	**0.854***	0.740
DSSM+VAEE	**0.831***	0.850	**0.827*****
NMF	0.794	0.869	0.810
NMF+VAE	0.803	0.869	0.822
NMF+VAEE	**0.806***	**0.871**	**0.826***
DeepFM	0.730	0.826	0.749
DeepFM+VAE	0.769***	0.870***	0.761*
DeepFM+VAEE	**0.769*****	**0.872*****	**0.765***

Note. The * denotes the significance levels of pairwise comparison t-tests between Baselines and our proposed Baseline+VAEs or Baseline+VAEEs. * p ⩽ 0.05, **p⩽ 0.01, ***p⩽ 0.001.

C. Explainability of Anchoring Effect

In this section, we provide more evidence to illustrate the existence of Anchoring Effect. To examine the effectiveness of the selected anchoring features, additional experiments are conducted to compare with models trained on different user feature subsets when learning the anchoring representation. As we expected, the anchoring related features (i.e. user-id and gender), which are summarized by previous research, should show significant difference on quantifying the anchoring effect from the other feature subsets.

Experiment results are displayed in Table III. Several conclusions can be made: 1) In most of the cases, more user features as the input of our VAEE model can bring about better performance on entire accuracy. 2) However, the subset of *user-id + gender* still outperform the other feature subsets on AUC. It can be concluded that the current anchoring feature subset is the most efficient one to fit our anchoring representation learning module. The result also provides experimental confirmation for the empirical conclusions on anchoring effect in former researches [19,22]. 3) The pairwise significance t-tests of final AUC results are conducted between using *user-id+gender* as input and other feature subsets as input. Comparing with other feature subsets, using *user-id+gender* as the anchoring feature input acquires significantly better performance in our prediction model. This also announces the importance of considering *user-id+gender* as anchoring features.

TABLE III. AUC OF OFFLINE EXPERIMENTS WITH DSSM + VAEE FOR DIFFERENT ANCHOR FEATURE SUBSETS ON AMAZON DATASET

Anchor Features	CIKM 2019 Ecomm AI
user-id	0.827
age	0.823***
gender	0.824*
user-id+gender	**0.831**
user-id+age	0.828**
user-id+gender+age	0.830**

Note. The * denotes the significance levels of pairwise comparison t-tests between *user-id+gender* as input and other feature subsets as input. * p ⩽ 0.05, **p⩽ 0.01, ***p⩽ 0.001.

V. CONCLUSIONS

This paper proposes a novel model based on VAE to mitigate the anchoring effect in recommendations. Experiments have provided evidence that the anchoring representation can be captured with our proposed VAEE model and reducing anchoring effect can help recommendation models perform better. The proposed VAEE is attachable to most recommendation models, as an additional learning framework to quantify and mitigate the anchoring effect, for better understanding of users' true preferences. Our limitation includes the selection of anchoring features. The background support of empirical researches still need to be examined in real-world online experiments. Future works are expected to test the model performance on various combinations of user features and reconstruct users' true preferences.

REFERENCES

[1] Zhang, S., Yao, L., Sun, A., and Tay, Y. 2019."Deep Learning Based Recommender System: A Survey and New Perspectives," ACM computing surveys (CSUR) (52:1), pp. 1–38.

[2] Guo, H., Tang, R., Ye, Y., Li, Z., and He, X. 2017."Deepfm: A Factorization-Machine Based Neural Network for CTR Prediction," arXiv preprint arXiv:1703, 04247.

[3] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T. 2017."Neural Collaborative Filtering," In Proceedings of the 26th international conference on world wide web, pp. 173–182.

[4] Tversky, A., and Kahneman, D. 1974.“Judgment under Uncertainty: Heuristics and Biases: Biases in Judgments Reveal Some Heuristics of Thinking under Uncertainty,” Science (185:4157), pp. 1124–1131.

[5] Adomavicius, G., Bockstedt, J. C., Curley, S. P., and Zhang, J. J. 2013.“Do Recommender Systems Manipulate Consumer Preferences? A Study of Anchoring Effects,” Information Systems Research (24:4), pp. 956–975.

[6] Cosley, D., Lam, S. K., Albert, I., Konstan, J. A., and Riedl, J. 2003.“Is Seeing Believing? How Recommender System Interfaces Affect Users' Opinions,” In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 585–592.

[7] Köcher, S., Jugovac, M., Jannach, D., and Holzmüller, H. H. 2019.“New Hidden Persuaders: An Investigation of Attribute-Level Anchoring Effects of Product Recommendations,” Journal of Retailing (95:1), pp. 24–41.

[8] Felfernig, A., Friedrich, G., Gula, B., Hitz, M., Kruggel, T., Leitner, G., Melcher, R., Riepan, D., Strauss, S., and Teppan, E. 2007.“Persuasive Recommendation: Serial Position Effects in Knowledge-Based Recommender Systems,” In Persuasive Technology: Second International Conference on Persuasive Technology, pp. 283–294.

[9] Zhao, Z., Chen, J., Zhou, S., He, X., Cao, X., Zhang, F., and Wu, W. 2022.“Popularity Bias Is Not Always Evil: Disentangling Benign and Harmful Bias for Recommendation,” IEEE Transactions on Knowledge and Data Engineering.

[10] Vardasbi, A., Oosterhuis, H., and Rijke, M. de 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, pp. 1475–1484.

[11] Lin, C., Liu, X., Xv, G., and Li, H. 2021.“Mitigating Sentiment Bias for Recommender Systems,” In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 31–40.

[12] Liang, D., Krishnan, R. G., Hoffman, M. D., and Jebara, T. 2018.“Variational Autoencoders for Collaborative Filtering,” In Proceedings of the 2018 world wide web conference, pp. 689–698.

[13] Ma, W., Chen, X., Pan, W., and Ming, Z. 2022.“Vae++ Variational Autoencoder for Heterogeneous One-Class Collaborative Filtering,” In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, pp. 666–674.

[14] Ding, Y., Shi, Y., Chen, B., Lin, C., Lu, H., Li, J., Tang, R., and Wang, D. 2021.“Semi-Deterministic and Contrastive Variational Graph Autoencoder for Recommendation,” In Proceedings of the 30th ACM International Conference on Information & Knowledge Management, pp. 382–391.

[15] Gao, Z., Shen, T., Mai, Z., Bouadjenek, M. R., Waller, I., Anderson, A., Bodkin, R., and Sanner, S. 2022.“Mitigating the Filter Bubble While Maintaining Relevance: Targeted Diversification With Vae-Based Recommender Systems,” In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 2524–2531.

[16] Nema, P., Karatzoglou, A., and Radlinski, F. 2021.“Disentangling Preference Representations for Recommendation Critiquing with ß-Vae,” In Proceedings of the 30th ACM International Conference on Information & Knowledge Management, pp. 1356–1365.

[17] Salah, A., Binh, T., and Lauw, H. 2021.“Towards Source-Aligned Variational Models for Cross-Domain Recommendation,” In Proceedings of the 15th ACM Conference on Recommender Systems, pp. 176–186.

[18] Huang, P., He, X., Gao, J., Deng, L., Acero, A., and Heck, L. 2013.“Learning Deep Structured Semantic Models for Web Search Using Clickthrough Data,” In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, pp. 2333–2338.

[19] Adomavicius, G., Bockstedt, J. C., Curley, S. P., and Zhang, J. J. 2018.“Effects of Online Recommendations on Consumers' Willingness to Pay,” Information Systems Research (29:1), pp. 84–102.

[20] Kingma, D. P., and Welling, M. 2013.“Auto-Encoding Variational Bayes,” arXiv preprint arXiv:1312, 6114.

[21] Epley, N., and Gilovich, T. 2005.“When Effortful Thinking Influences Judgmental Anchoring: Differential Effects of Forewarning and Incentives on Self-Generated and Externally Provided Anchors,” Journal of Behavioral Decision Making (18:3), 199–212.

[22] Adomavicius, G., Bockstedt, J. C., Curley, S. P., and Zhang, J. J. 2019.“Reducing Recommender System Biases: An Investigation of Rating Display Designs,” MIS Quarterly (43:4), pp. 1321–1341.

[23] Kahneman, D., Rosenfield, A., Gandhi, L., and Blaser, T. 2016.“Noise,” Harvard business review (94), pp. 38–46.

[24] Lee, C., and Morewedge, C. K. 2022.“Noise Increases Anchoring Effects,” Psychological Science (33:1), pp. 60–75.

[25] Wan, M., Ni, J., Misra, R., and McAuley, J. 2020.“Addressing Marketing Bias in Product Recommendations,” In Proceedings of the 13th international conference on web search and data mining, pp. 618–626.

[26] Harper, F. M., Konstan, and J. A. 2015.“The Movielens Datasets: History and Context,” Acm transactions on interactive intelligent systems (TIIS) (5:4), pp. 1–19.

[27] Kingma, D. P., and Ba, J. 2014.“Adam: A Method for Stochastic Optimization,” arXiv preprint arXiv:1412, 6980.