The Fortieth AAAI Conference on Artificial Intelligence (AAAI-26)

Bidirectional Counterfactual Distillation for Review-Based Recommendation

Sheng Sang[1], **Shujie Li**[2], **Shuaiyang Li**[3], **Kang Liu**[4], **Teng Li**[5], **Wei Jia**[1], **Dan Guo**[1], **Feng Xue**[2]*
[1]School of Computer Science and Information Engineering, Hefei University of Technology
[2]School of Software, Hefei University of Technology
[3]College of Computer and Information Engineering, Henan Normal University
[4]School of Health Management, Anhui Medical University
[5]School of Artificial Intelligence, Anhui University
meetssang@gmail.com, lisjhfut@hfut.edu.cn, lishuaiyang@htu.edu.cn, kangliu1225@gmail.com, liteng@ahu.edu.cn,
jiawei@hfut.edu.cn, guodan@hfut.edu.cn, feng.xue@hfut.edu.cn

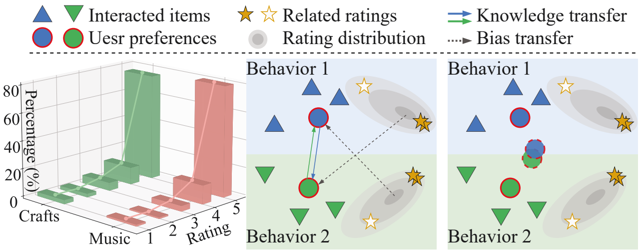
Abstract

Review-based recommendation methods typically integrate multiple behaviors, including interactions, reviews, and ratings, to model user preferences. To effectively extract preference signals from diverse behaviors, some studies train multiple student models to capture distinct behavioral patterns, and leverage online distillation to facilitate collaborative learning among them. However, we argue that these techniques suffer from *bias contamination from rating distributions* and *feature homogenization* during cross-behavior knowledge transfer: (1) Rating distribution bias, arising from non-uniform historical ratings, propagates across behaviors through distillation, contaminating the true preference representations of other behaviors. (2) Static distillation strategies often lead to homogenized behavioral features, hindering the learning of behavior-specific preferences. To address these issues, we propose a novel Bidirectional Counterfactual Distillation (BiCoD) framework for review-based recommendation. In BiCoD, we first design an adversarial counterfactual distillation module to suppress the impact of non-uniform rating distributions on distillation, thereby preventing it from contaminating the user's true preference representations across behaviors. Subsequently, we introduce a stage-aware bidirectional distillation strategy to enhance the distinctiveness of behavioral features, facilitating the effective learning of behavior-specific preferences. Extensive experiments on five real-world datasets validate the effectiveness and superiority of the proposed framework.

Code — https://github.com/hfutmars/BiCoD

Introduction

As a widely adopted recommendation technique, review-based methods (Liu et al. 2021; Shuai et al. 2022; Ren et al. 2023) typically leverage users' historical behavioral data (e.g., past interactions, reviews, and ratings) to capture user preferences and predict ratings for target items (Chen et al. 2018; Li et al. 2021; Wu et al. 2023). Recently, most review-based methods (Shuai et al. 2022; Ren et al. 2023) construct a unified graph to integrate various types of user behaviors, where ratings are represented as edges, and adopt message passing to aggregate features from different behaviors.



(a) Rating distribution (b) Bias contamination (c) Feature homogenization

Figure 1: Motivation example. (a) Real-world rating distributions are typically non-uniform. (b) Online distillation updates user preferences based on both cross-behavior signals (solid) and rating distribution bias (dashed). (c) Static distillation leads to feature homogenization, limiting behavior-specific preference learning.

However, different behaviors inherently convey distinct levels of user intent (e.g., interactions represent general interests, reviews reflect subjective feedback, and aspects reveal fine-grained preferences (Li et al. 2021)). Treating all behaviors uniformly in a single graph may lead the model to overlook these semantic differences when aggregating neighbor features. Therefore, some studies (Du et al. 2023; Zhu and Zhang 2025) train multiple student models in parallel to capture distinct behavioral features, and adopt feature-based online distillation (Yang et al. 2023a) to facilitate knowledge transfer among them, thereby improving rating prediction accuracy. Despite these advances, existing online distillation methods indiscriminately transfer all knowledge from each student, leading to the following issues:

1) **Bias contamination from rating distributions**: Due to the typically non-uniform distribution of historical ratings, high-frequency ratings often dominate the learning process, causing models to tend to overfit the rating distribution patterns to minimize prediction loss rather than accurately modeling users' true preferences. We refer to this tendency as ***rating distribution bias***. Although some methods (Xv et al. 2022; Du et al. 2023) attempt to alleviate such

---

*Corresponding author.

bias in individual behaviors by employing multiple teacher models, they overlook the fact that this bias can also propagate during online knowledge distillation, thereby contaminating the preference representations of other behaviors. As shown in Figure 1(a) and 1(b), user preference updates are influenced by the rating distributions of other behaviors.

2) **Feature homogenization among behaviors**: Existing online feature distillation strategies adopt fixed distillation directions or strengths, which aligns student models' knowledge into a shared subspace. However, such paradigms often lead to the gradual homogenization of feature representations across different behavioral branches, limiting the student model's ability to learn the distinct intentions inherent in each behavior. As shown in Figure 1(c), the preference representations of different behaviors gradually converge.

To address the aforementioned issues, we argue that it is necessary to reorganize both the knowledge and strategy of distillation to tackle the complexities of multi-behavior feature learning in review-based recommendation. Specifically, we identify two core challenges:

1) How to effectively mitigate the rating distribution bias in each student model's knowledge during distillation, prevent its cross-behavior propagation, and adapt to the continuous evolution of the student models?

2) How to reformulate the distillation strategy to facilitate knowledge transfer across behaviors while preserving the distinctiveness of each behavior's feature representation?

To this end, we propose a novel **Bidirectional Counterfactual Distillation** (BiCoD) framework for review-based recommendation, which aims to address bias contamination and feature homogenization in multi-behavior online feature distillation:

First, we design an ***adversarial counterfactual distillation*** module to tackle the first challenge. To suppress the influence of non-uniform rating distributions on online distillation, we leverage *counterfactual reasoning* to estimate the causal discrepancy between the factual user-item bipartite graph and its counterfactual graph that retains only rating signals, thereby purifying the behavioral representations learned by each student model. Furthermore, *adversarial learning* is employed to dynamically adjust the strength of bias suppression, thus adapting to the evolving student models and balancing online distillation with bias suppression.

Subsequently, we introduce a ***stage-aware bidirectional distillation*** strategy to tackle the second challenge. In the early stages of training, *forward distillation* is applied to promote knowledge transfer across behaviors, enhancing the representation capacity of each student model. In later stages, *reverse distillation* is employed to increase the discrepancy between behavioral representations, thereby enabling the model to capture behavior-specific preferences.

The key contributions are summarized as follows:

- This work highlights the issues of bias contamination and feature homogenization in online distillation paradigms, and theoretically analyzes their underlying causes.
- We design a counterfactual distillation module to prevent the propagation of rating distribution bias during distillation, and employ adversarial learning to adapt to the continuously evolving student model.
- We propose a stage-aware bidirectional distillation strategy to enhance knowledge transfer across behaviors while capturing the distinct features of each behavior.
- Extensive experiments on five real-world datasets validate the effectiveness and superiority of BiCoD.

Related Work

Existing review-based recommendation methods mainly adopt dual-tower architectures (Chen et al. 2018) and graph neural networks to model multi-source behaviors (Xia, Huang, and Zhang 2022; Yang et al. 2023b). Other methods (Bauman, Liu, and Tuzhilin 2017; Li et al. 2021) extract aspect-level signals from reviews to model both implicit textual features and explicit aspect information. Recent methods unify diverse behaviors into a graph to enable cross-behavior information aggregation (Shuai et al. 2022; Ren et al. 2023; Wang et al. 2024). For instance, RGCL (Shuai et al. 2022) incorporates review embeddings as graph edges to capture fine-grained user and item representations.

Knowledge distillation (KD) (Zhang et al. 2018; Wu and Gong 2021; Kweon, Kang, and Yu 2021; Houyon et al. 2023; Ji et al. 2023) improves the performance of the target model by transferring knowledge across models (Guo et al. 2020; Lan, Zhu, and Gong 2018; Ding et al. 2022; Liu et al. 2023; Gou et al. 2024). For example, KDCRec (Liu et al. 2022) distills knowledge from uniform data to correct biases caused by non-uniform training. FreqD (Zhu and Zhang 2025) enhances critical knowledge by reweighting frequency components of feature representations. However, these methods overlook the issues of bias contamination and feature homogenization during knowledge distillation.

Preliminaries

Problem Definition

In review-based recommendation, there are two sets of entities: the user set $\mathcal{U}$ and the item set $\mathcal{V}$, where user $u \in \mathcal{U}$ and item $v \in \mathcal{V}$. Let $\mathcal{E} = \mathcal{U} \cup \mathcal{V}$ denote the set of all entities, where $e \in \mathcal{E}$ denotes an entity. The review text and rating given by user u to item v are denoted as t_{uv} and r_{uv}, respectively. Let $\mathcal{T}$ denote the set of all reviews. The set of all possible ratings is denoted by $\mathcal{R} = \{r_1, \dots, r_{|\mathcal{R}|}\}$ (e.g., $\mathcal{R} = \{1, \dots, 5\}$ in Amazon), where $|\cdot|$ denotes the cardinality of a set. The goal of review-based recommendation is to predict the rating $\hat{r}_{uv}$ that user u would assign to item v, by leveraging their historical interactions, reviews, and ratings.

Theoretical Analysis of Online Distillation

In this section, we briefly analyze how the rating distribution bias in one behavior contaminates the true preference representation of another from the perspective of gradients. For any two behaviors x and y, the gradient of the SCE loss (Hou et al. 2022) with respect to $\boldsymbol{e}_y$ is given by:

$$\nabla_{\boldsymbol{e}_y} f_{\text{sce}}(\boldsymbol{e}_x, \boldsymbol{e}_y) = -\frac{\boldsymbol{e}_x}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|} + \frac{\boldsymbol{e}_x^\top \boldsymbol{e}_y}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|^3}\, \boldsymbol{e}_y. \quad (1)$$

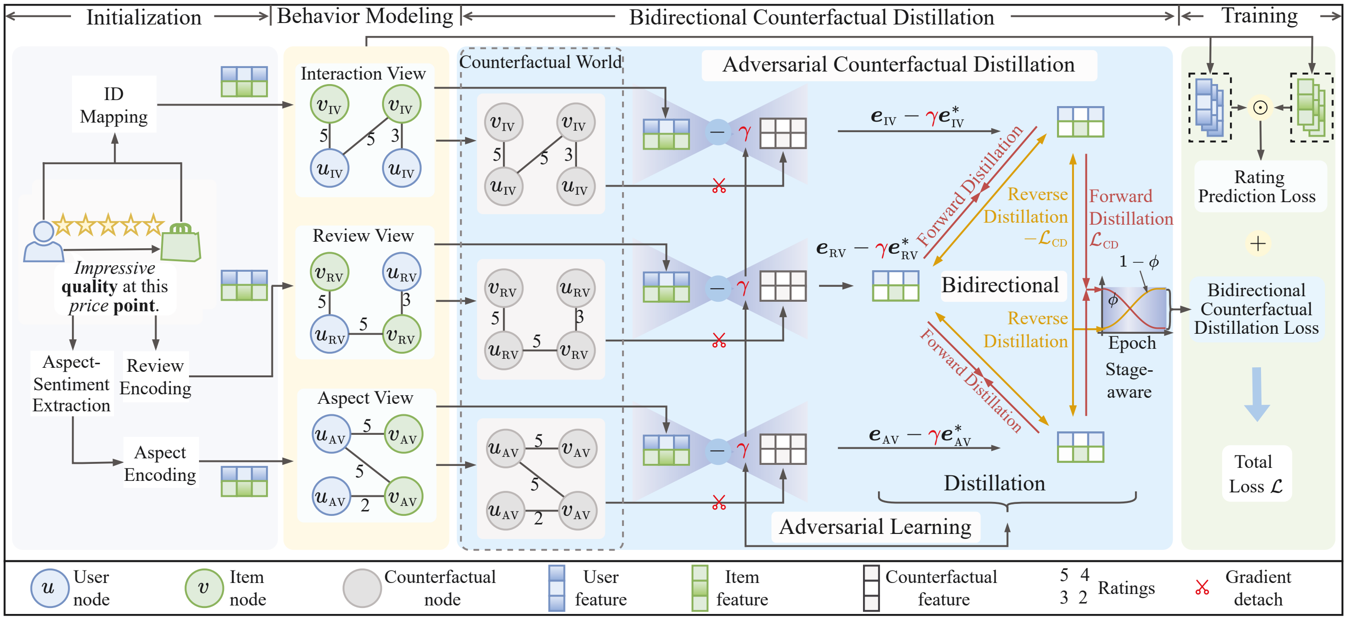


Figure 2: The architecture of the bidirectional counterfactual distillation framework.

where $f_{\text{sce}}(\boldsymbol{e}_x, \boldsymbol{e}_y) = 1 - \frac{\boldsymbol{e}_x^\top \boldsymbol{e}_y}{\|\boldsymbol{e}_x\|\cdot\|\boldsymbol{e}_y\|}$, and $\boldsymbol{e}_x$ and $\boldsymbol{e}_y$ denote the entity features from different behaviors.

We decompose each representation into a true preference component (TIE) and a bias component (NDE), analogous to causal-effect decomposition (Wei et al. 2021; Li et al. 2023): $\boldsymbol{e}_x = \text{TIE}_x + \text{NDE}_x, \quad \boldsymbol{e}_y = \text{TIE}_y + \text{NDE}_y$, where TIE (total indirect effect) represents the ideal preference signal, and NDE (natural direct effect) denotes the bias component in each view.

The first term can be rewritten as

$$-\frac{\boldsymbol{e}_x}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|} = -\frac{\text{TIE}_x + \text{NDE}_x}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|}. \tag{2}$$

It indicates that the bias component NDE_x in view x directly influences the update direction of features in view y, thereby causing NDE_x to contaminate TIE_y.

The second term $\frac{\boldsymbol{e}_x^\top \boldsymbol{e}_y}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|^3}\,\boldsymbol{e}_y$ can be expanded as

$$\frac{(\text{TIE}_x + \text{NDE}_x)^\top(\text{TIE}_y + \text{NDE}_y)}{\|\boldsymbol{e}_x\|\|\boldsymbol{e}_y\|^3} \cdot (\text{TIE}_y + \text{NDE}_y). \tag{3}$$

The inner product expands to four components: $\text{TIE}_x^\top\text{TIE}_y$, which represents the desired knowledge alignment; $\text{TIE}_x^\top\text{NDE}_y$ and $\text{NDE}_x^\top\text{TIE}_y$, which indicate cross-behavior bias contamination; and $\text{NDE}_x^\top\text{NDE}_y$, which reflects bias reinforcement. Furthermore, the multiplicative term $(\text{TIE}_y + \text{NDE}_y)$ further amplifies the self-propagation of NDE_y, allowing the bias in $\boldsymbol{e}_y$ to contaminate behavior representations through gradient flow.

Therefore, eliminating the bias components from each behavior during distillation is essential to ensure that the distillation loss aligns only with the true preferences, thereby preventing cross-behavioral bias contamination.

Proposed Framework

Figure 2 illustrates the overall architecture of BiCoD. First, multiple parallel student networks are constructed to capture various behavioral patterns, including interactions, reviews, and aspect-sentiment behaviors. Next, ***adversarial counterfactual distillation*** is employed to suppress the contamination of true user preferences by rating distribution bias during distillation. Then, a ***stage-aware bidirectional distillation*** strategy is introduced to facilitate knowledge sharing while avoiding feature homogenization. Finally, the model is optimized by jointly minimizing the rating prediction loss and the distillation loss.

Initialization

The representation of an entity consists of the following features: ID features, review semantic features, aspect-sentiment features from reviews, and rating features.

For the user and item IDs, we map them into dense vectors, denoted as $\boldsymbol{e}_{\text{id},u} \in \mathbb{R}^d$ and $\boldsymbol{e}_{\text{id},v} \in \mathbb{R}^d$, where d represents the embedding dimension. Each review in the set $\mathcal{T}$ is first fed into a pre-trained BERT-Whitening (Su et al. 2021) model to obtain sentence embeddings. Then, the review embeddings associated with each user and item are averaged to obtain the review features $\boldsymbol{e}_{\text{re},u} \in \mathbb{R}^d$ and $\boldsymbol{e}_{\text{re},v} \in \mathbb{R}^d$, respectively. For aspect-sentiment features, we first extract aspect-sentiment pairs from each review using the SPAN-aste (Xu, Chia, and Bing 2021) model. These pairs are then encoded into embeddings via BERT-Whitening, and subsequently averaged to obtain $\boldsymbol{e}_{\text{as},u} \in \mathbb{R}^d$ and $\boldsymbol{e}_{\text{as},v} \in \mathbb{R}^d$. For rating features, all discrete ratings in $\mathcal{R}$ are mapped into a continuous embedding space. The rating feature corresponding to a specific user-item pair (u, v) is denoted as $\boldsymbol{r}_{uv}$.

Therefore, the review-based recommendation data can be formulated as a bipartite graph $\mathcal{G} = \langle \mathcal{U} \cup \mathcal{V}, \mathcal{R} \rangle$. Each observed edge corresponds to a rating embedding from user u to item v, which directly reflects the user's overall satisfaction with the item.

Multi-Behavior Modeling (MB)

To comprehensively capture user interests and item features under different behavioral patterns, we construct three parallel views to model interaction behavior, review behavior, and aspect-sentiment behavior, respectively:

- Interaction View (IV): models interaction behavior, where node features are user and item ID embeddings.
- Review View (RV): captures review behavior, where node features are derived from the review representations of users and items.
- Aspect View (AV): models the aspect-sentiment behavior, where nodes represent aspect-sentiment features for both users and items.

For each view, edges represent rating embeddings within user-item bipartite graph. We employ graph convolutional networks (He et al. 2020; Shuai et al. 2022) to aggregate information from source to target nodes:

$$\boldsymbol{e}_{m,u}^{(l)} = \sum_{v \in \mathcal{N}_u} \frac{\boldsymbol{e}_{m,v}^{(l-1)} \boldsymbol{r}_{m,uv}}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_v|}}, \tag{4}$$

where $m \in \mathcal{M}$ denotes a specific view, and $\mathcal{M} = \{\text{IV}, \text{RV}, \text{AV}\}$ represents the set of all views. In view m, $\boldsymbol{e}_{m,u}^{(l)}$ and $\boldsymbol{e}_{m,v}^{(l-1)}$ denote the l-th layer embedding of user u and the $(l-1)$-th layer embedding of item v, respectively. $\boldsymbol{r}_{m,uv}$ denotes the embedding of the rating given by user u to item v. $\mathcal{N}_u$ denotes the set of items that user u has rated, and $\mathcal{N}_v$ indicates the set of users who have rated item v. When $(l = 0)$, $\boldsymbol{e}_{\text{IV},v}^{(0)} = \boldsymbol{e}_{\text{id},v}$, $\boldsymbol{e}_{\text{RV},v}^{(0)} = \boldsymbol{w}_1^\top \boldsymbol{e}_{\text{re},v}$, and $\boldsymbol{e}_{\text{AV},v}^{(0)} = \boldsymbol{w}_2^\top \boldsymbol{e}_{\text{as},v}$, where $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathbb{R}^d$ are learnable parameters, and $\top$ denotes the transpose operation.

Similarly, the item representation $\boldsymbol{e}_{\text{m},v}^{(l)}$ under view m at layer l can also be obtained through graph convolution operations. We stack L graph convolutional layers and use the final outputs, $\boldsymbol{e}_{m,u}^{(L)}$ and $\boldsymbol{e}_{m,v}^{(L)}$, as the ultimate user and item features, denoted by $\boldsymbol{e}_{m,u}$ and $\boldsymbol{e}_{m,v}$, respectively.

Adversarial Counterfactual Distillation

Online Distillation (OD) To enhance the feature representations of entities (users or items) within different behaviors, we leverage feature-based online distillation to promote knowledge transfer for the same entity across multiple behavioral views. For the online distillation task, the MB module is optimized by jointly minimizing the rating prediction loss and the distillation loss. The objective of the conventional distillation task is defined as follows:

$$\min_{\theta_{\text{MB}}} \mathcal{L}_{\text{OD}}(\theta_{\text{MB}}) = \min_{\theta_{\text{MB}}} \frac{1}{|\mathcal{E}|} \sum_{k \in \mathcal{E}} \sum_{x \in \mathcal{M}; x \neq y} \sum_{y \in \mathcal{M}} f_{\text{sce}}(\boldsymbol{e}_{x,k}, \boldsymbol{e}_{y,k}), \tag{5}$$

where θ_{MB} denotes the learnable parameters of the MB module, and $\boldsymbol{e}_{x,k}, \boldsymbol{e}_{y,k} \in \mathbb{R}^d$ are the feature representations of entity $k \in \mathcal{E} = \mathcal{U} \cup \mathcal{V}$ in views x and y, respectively. $f_{\text{sce}}(x, y) = (1 - \frac{x^\top y}{\|x\| \cdot \|y\|})^\delta$ denotes the Scaled Cosine Error (SCE) loss (Hou et al. 2022), which is adopted to align the feature representations of the same entity k across views x and y. The scaling factor δ of SCE is fixed to 1 in this paper.

Counterfactual Distillation (CD) To prevent rating distribution bias from contaminating user preferences across behaviors through distillation, it is necessary to eliminate such bias from entity features, thereby enabling online distillation to align only the true preference representations among behaviors. Motivated by causal inference (Wei et al. 2021; Li et al. 2023), the causal discrepancy between factual and counterfactual representations can quantify the actual effect of user preferences on predicted ratings. Accordingly, we first explore the following counterfactual question to obtain the entity features of the counterfactual world: What would the features of the user and item be if they were determined solely by historical ratings? In this counterfactual world, the graph aggregation process transforms into:

$$\boldsymbol{e}_{m,u}^{(l)^*} = \sum_{v \in \mathcal{N}_u} \frac{\boldsymbol{r}_{m,uv}}{\sqrt{|\mathcal{N}_u|}\sqrt{|\mathcal{N}_v|}}, \tag{6}$$

where $\boldsymbol{e}_{m,u}^{(l)^*}$ denotes the user feature learned in the counterfactual world. Similarly, the item features $\boldsymbol{e}_{m,v}^{(l)^*}$ can be computed under the same setting. The counterfactual learning process is detached from gradient backpropagation, as our goal is to train a model for the real world rather than the counterfactual one.

Subsequently, to mitigate the impact of rating distribution bias on distillation, we purify preference-related representations by subtracting counterfactual entity features from factual entity features during distillation:

$$\begin{aligned} \min_{\theta_{\text{MB}}} \mathcal{L}_{\text{CD}}(\theta_{\text{MB}}) = \min_{\theta_{\text{MB}}} \frac{1}{|\mathcal{E}|} \sum_{k \in \mathcal{E}} \sum_{x \in \mathcal{M}; x \neq y} \\ \sum_{y \in \mathcal{M}} f_{\text{sce}}(\boldsymbol{e}_{x,k} - \gamma \boldsymbol{e}_{x,k}^*, \boldsymbol{e}_{y,k} - \gamma \boldsymbol{e}_{y,k}^*), \end{aligned} \tag{7}$$

where $\gamma \in \mathbb{R}$ is a learnable parameter that controls the strength of bias suppression. Notably, when γ is fixed at 0, the objective degenerates into conventional online distillation. A larger γ indicates a greater removal of bias, but retaining moderate bias can be beneficial (Qiu et al. 2021). Moreover, the optimal γ may vary across datasets and training stages, making a fixed γ insufficient to accommodate the bias of the evolving student model.

Adversarial Learning (AL) For models trained on biased real-world data, the gradient of the counterfactual distillation loss with respect to γ satisfies $\frac{\partial \mathcal{L}_{\text{CD}}}{\partial \gamma} > 0$. During gradient descent optimization, γ is updated as:

$$\gamma \leftarrow \gamma - \eta \frac{\partial \mathcal{L}_{\text{CD}}}{\partial \gamma}, \tag{8}$$

where η denotes the learning rate. This update causes γ to gradually decrease, preventing the model from effectively mitigating the propagation of rating distribution bias during distillation. *In other words, models trained on real-world data with non-uniform rating distributions naturally tend to learn the rating distribution patterns to minimize training loss.* However, this optimization trend conflicts with our goal of suppressing bias propagation during distillation, which requires encouraging a larger value of γ.

Therefore, we introduce an adversarial learning approach (Goodfellow et al. 2014) to dynamically adjust the strength γ to adapt to the evolving student model, thereby effectively suppressing the influence of the rating distribution on distillation. Specifically, while the MB module seeks to minimize $\mathcal{L}_{\text{CD}}$ by fitting the rating distribution, the parameter γ is adversarially optimized to maximize $\mathcal{L}_{\text{CD}}$, thereby suppressing bias contamination from rating distributions. Unlike conventional distillation, the MB module and γ form a two-player mini-max game, with the value function defined as:

$$\min_{\theta_{\text{MB}}} \max_{\theta_{\text{AL}}} \mathcal{L}_{\text{CD}}(\theta_{\text{MB}}) = \min_{\theta_{\text{MB}}} \max_{\theta_{\text{AL}}} \frac{1}{|\mathcal{E}|} \sum_{k \in \mathcal{E}} \sum_{x \in \mathcal{M}; x \neq y} \sum_{y \in \mathcal{M}} f_{\text{sce}}(\boldsymbol{e}_{x,k} - \gamma \boldsymbol{e}^*_{x,k}, \boldsymbol{e}_{y,k} - \gamma \boldsymbol{e}^*_{y,k}), \tag{9}$$

where θ_{AL} denotes the parameters updated via adversarial learning, i.e., γ.

An alternating algorithm is employed to address the problem in Equation 9, in which one set of variables is fixed while the other is optimized at each iteration:

$$\begin{aligned} \hat{\theta}_{\text{MB}} &= \arg \min_{\theta_{\text{MB}}} \mathcal{L}(\theta_{\text{MB}}, \hat{\theta}_{\text{AL}}), \\ \hat{\theta}_{\text{AL}} &= \arg \max_{\theta_{\text{AL}}} \mathcal{L}(\hat{\theta}_{\text{MB}}, \theta_{\text{AL}}), \end{aligned} \tag{10}$$

where $\mathcal{L}$ represents the total loss of the model.

In practice, we implement the above adversarial process using a Gradient Reversal Layer (GRL) (Ganin and Lempitsky 2015). The model parameters are updated as follows:

$$\begin{aligned} \theta_{\text{MB}} &\leftarrow \theta_{\text{MB}} - \eta \frac{\partial \mathcal{L}}{\partial \theta_{\text{MB}}}, \\ \theta_{\text{AL}} &\leftarrow \theta_{\text{AL}} + \eta \frac{\partial \mathcal{L}}{\partial \theta_{\text{AL}}}. \end{aligned} \tag{11}$$

Stage-aware Bidirectional Distillation

For online distillation, adopting a fixed direction of knowledge transfer throughout all training stages may be suboptimal. Analogous to real-world education, excessive knowledge sharing leads to over-homogenization among students, diminishing their individuality and reducing their ability to cope with diverse social environments. Humans tend to learn more effectively when tasks are organized in a meaningful order (Bengio et al. 2009; Zheng et al. 2021).

Accordingly, we introduce a stage-aware curriculum approach that dynamically scales the distillation loss $\mathcal{L}_{\text{CD}}$. At early training stages, entity representations of different behavioral views are unstable. Therefore, we adopt Forward Distillation (FD) to extract additional supervisory signals from other behaviors, guiding each view toward shared feature alignment and improving the representation quality. As training progresses, entity representations gradually converge. To avoid excessive homogenization, we progressively weaken Forward Distillation and enhance Reverse Distillation (RD) to increase the discrepancy among behaviors, guiding the model to focus on behavior-specific representations and better capture diverse behavioral patterns.

Formally, we dynamically adjust the distillation direction and strength according to the following cosine schedule:

$$\mathcal{L}_{\text{BiCoD}} = \begin{cases} \phi_n \, \mathcal{L}_{\text{CD}}, & \text{if } E_n \bmod 2 = 1 \quad \text{(FD)} \\ -\big(1 - \phi_n\big) \, \mathcal{L}_{\text{CD}}, & \text{otherwise} \quad \text{(RD)} \end{cases} \tag{12}$$

where E_n denotes the n-th epoch. E_m denotes the maximum number of epochs. The scheduling coefficient $\phi_n = \max\Big\{0.1,\ \frac{1}{2}\Big[\cos\Big(\frac{\min(E_n, E_m)}{E_m}\pi\Big) + 1\Big]\Big\}$ varies with the epoch, following a cosine decay from 1 to 0.1.

This scheduling mechanism enables stage-aware adjustment of both the direction and strength of online distillation without adding extra model parameters. It effectively mitigates representation homogenization and enhances the learning of more distinctive behavior-specific patterns.

Training and Inference

Training We first compute user–item matching features, then predict the rating via a Multi-Layer Perceptron (MLP) with a non-linear transformation:

$$\hat{r}_{m,uv} = \mathcal{R}^\top \cdot \text{Softmax}(\text{MLP}_m(\boldsymbol{e}_{m,u} \odot \boldsymbol{e}_{m,v})), \tag{13}$$

where $\hat{r}_{m,uv}$ denotes the predicted rating between user u and item v in view m, $\odot$ is the Hadamard product, and $\text{MLP}_m(\cdot)$ is a two-hidden-layer network with LeakyReLU (negative slope 0.1) activation functions.

Rating prediction is treated as a regression task (Shuai et al. 2022) and optimized with Mean Squared Error (MSE):

$$\min_{\theta_{\text{MB}}} \mathcal{L}_{\text{RP}}(\theta_{\text{MB}}) = \min_{\theta_{\text{MB}}} \Big(\frac{1}{|\mathcal{S}|} \sum_{(u,v) \in \mathcal{S}} \sum_{m \in \mathcal{M}} (\hat{r}_{m,uv} - r_{uv})^2 \Big), \tag{14}$$

where $\mathcal{L}_{\text{RP}}$ denotes the loss for rating prediction. $\mathcal{S}$ represents user-item pairs in the training set.

The overall training objective is formulated as follows:

$$\min_{\theta_{\text{MB}}} \max_{\theta_{\text{AL}}} \mathcal{L}(\theta_{\text{MB}}, \theta_{\text{AL}}) = \min_{\theta_{\text{MB}}} \max_{\theta_{\text{AL}}} \big(\mathcal{L}_{\text{RP}}(\theta_{\text{MB}}) + \mathcal{L}_{\text{BiCoD}}(\theta_{\text{MB}}, \theta_{\text{AL}}) \big). \tag{15}$$

Inference We compute the average of predictions across all views as the final predicted rating:

$$\hat{r}_{uv} = \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \hat{r}_{m,uv}. \tag{16}$$

Experiments

Experimental Setup

Dataset: We evaluate our model on five real-world datasets: Art_Crafts, Digital_Music, Luxury_Beauty, and Software from the Amazon review corpus, and the Yelp review

Models	Art_Crafts		Digital_Music		Luxury_Beauty		Software		Yelp	
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
DeepCoNN	0.7855	0.5530	0.3080	0.2681	0.6765	0.6078	1.0589	0.7883	1.1861	0.8707
NARRE	0.7295	0.5460	0.2922	0.2594	0.7495	0.6011	1.1072	0.7912	1.1823	0.8536
DAML	0.7083	0.5394	0.2977	0.2615	0.7064	0.5875	1.0835	0.7803	1.1815	0.8610
APRE	0.7071	0.5337	0.2877	0.2406	0.6794	0.5710	1.1651	0.7983	1.1744	0.8501
RGCL	0.6629	0.5143	0.2829	0.2413	0.6715	0.5681	1.0740	0.7782	1.1644	0.8495
SGDN	0.6612	0.5102	0.2809	0.2542	0.6853	0.5626	1.0868	0.7768	1.1635	0.8449
MAGCL	0.6556	0.4966	0.2778	0.2112	0.6703	0.5573	1.0668	0.7705	1.1763	0.8534
ReHCL	0.6511	0.4961	0.2864	0.2344	0.6713	0.5337	1.0595	0.7713	1.1742	0.8579
KDCRec	0.6637	0.5294	0.2856	0.2345	0.6735	0.5329	1.0670	0.7831	1.1673	0.8511
LUME	0.6714	0.5377	0.2973	0.2351	0.6831	0.5334	1.0732	0.7857	1.1623	0.8567
EMKD	0.6532	0.5139	0.2764	0.2236	0.6754	0.5313	1.0601	0.7824	1.1596	0.8436
FreqD	0.6539	0.4983	0.2753	0.2097	0.6713	0.5345	1.0533	0.7941	1.1611	0.8494
BiCoD	**0.5963** *	**0.4572** *	**0.2403** *	**0.1863** *	**0.6237** *	**0.5065** *	**0.9798** *	**0.7405** *	**1.1193** *	**0.8167** *
% Imp.	8.42%	7.84%	12.71%	11.16%	6.95%	4.67%	6.98%	3.89%	3.48%	3.19%

Table 1: Overall performance comparison. Best results per column and baseline are bolded and underlined, respectively. $*$ indicates statistical significance based on a t-test (p-value < 0.01) compared to the second-best result. We report the average results over five runs for each experiment, along with the relative improvement (*% Imp.*) over the best-performing baseline.

dataset. Following previous studies (Shuai et al. 2022; Ren et al. 2023; Wang et al. 2023), each dataset is randomly split into training (80%), validation (10%), and test (10%) sets.

Evaluation Metrics: As a regression task, rating prediction aims to predict continuous rating values, rather than discrete class labels as in classification tasks. Therefore, we adopt MSE and MAE as evaluation metrics during testing, following prior studies (Shuai et al. 2022; Ren et al. 2023).

Baselines: We compare BiCoD with the following baselines, categorized into three groups: **Dual-tower architectures**: DeepCoNN (Zheng, Noroozi, and Yu 2017), NARRE (Chen et al. 2018), DAML (Liu et al. 2019), and APRE (Li et al. 2021); **GNN-based methods**: RGCL (Shuai et al. 2022), SGDN (Ren et al. 2023), MAGCL (Wang et al. 2023), and ReHCL (Wang et al. 2024); **KD-based methods**: KDCRec (Liu et al. 2022), LUME (Xv et al. 2022), EMKD (Du et al. 2023), and FreqD (Zhu and Zhang 2025).

Implementation Details: For all methods, the embedding dimension d is set to 64. Trainable parameters are initialized using Xavier initialization (Glorot and Bengio 2010). The parameter γ is initialized to 0.5. The graph convolutional layer depth L is set to 1. We adopt the Adam optimizer (Kingma and Ba 2017) with a learning rate η of 0.01. Early stopping is applied with a patience of 50 epochs, and the maximum number of training epochs (E_m) is set to 200.

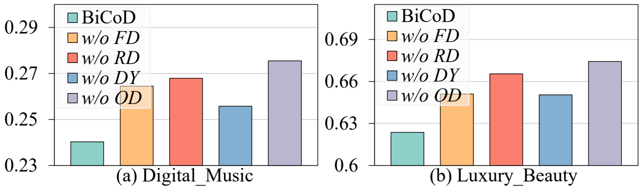
Overall Comparison

The performance of all methods is reported in Table 1. Our observations include: (1) Dual-tower methods (DeepCoNN, NARRE, DAML, and APRE) perform well, showing the value of incorporating reviews into recommendations. APRE's superior results highlight the benefit of using aspect-level review information. (2) GNN-based methods (RGCL, SGDN, MAGCL, and ReHCL) generally outperform dual-tower models, as they explicitly model the interactions among users, items, reviews, and aspects via graph neural networks. (3) Among distillation-based methods, KDCRec employs uniform sampling but lacks cross-behavior collaboration. LUME adopts multiple teachers yet misses unbiased knowledge sharing. EMKD neglects behavior-specific independence and rating distribution effects. FreqD reweights frequency components but still overlooks cross-behavior bias and feature homogenization. (4) Overall, BiCoD outperforms all baselines due to its adversarial counterfactual distillation that suppresses the influence of non-uniform rating distributions during distillation, and its bidirectional distillation that promotes knowledge transfer and behavioral distinctiveness.

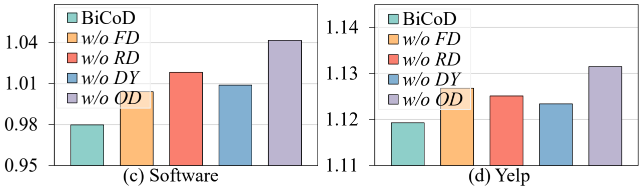
Ablation Study

Impact of Stage-aware Bidirectional Distillation. We compare BiCoD with the following variants to evaluate the effect of bidirectional distillation: *w/o FD*: removes forward distillation; *w/o RD*: removes reverse distillation; *w/o DY*: disables dynamic control by fixing the distillation weights of both FD and RD to 1; *w/o OD*: removes online distillation. Results are shown in Figure 3. Both *w/o FD* and *w/o RD* underperform BiCoD. This is because *w/o RD* leads to homogeneous entity representations, while *w/o FD* fails to leverage knowledge from other behaviors to learn representations for the target behavior. The performance drop of *w/o DY* stems from overly strong reverse distillation early on, hindering cross-behavior knowledge sharing, and excessive forward distillation later, which reduces the distinctiveness of behavior-specific features. The *w/o OD* variant performs the worst, as it lacks the mechanism for cross-behavior knowledge sharing, thus limiting feature enhancement.

Impact of Adversarial Counterfactual Distillation. We set up the following variants to evaluate adversarial counterfactual distillation: *w/o AL (FD/RD)* and *w/o CD (FD/RD)*



(a) Digital_Music (b) Luxury_Beauty



(c) Software (d) Yelp

Figure 3: MSE comparison of different variants.

Models	Art	Music	Beauty	Software	Yelp
w/o AL (FD)	0.6133	0.2461	0.6271	0.9905	1.1203
w/o CD (FD)	0.6211	0.2495	0.6492	0.9987	1.1210
w/o AL (RD)	0.6094	0.2433	0.6277	0.9802	1.1144
w/o CD (RD)	0.6159	0.2462	0.6463	0.9871	1.1207
w/o AL	0.6254	0.2489	0.6316	0.9959	1.1234
w/o CD	0.6306	0.2620	0.6561	1.0015	1.1246
w/ CD$_{\gamma=0.5}$	0.6251	0.2601	0.6337	0.9967	1.1236
w/ CD$_{\gamma=1}$	0.6195	0.2481	0.6277	1.0007	1.1257
BiCoD	**0.5963**	**0.2403**	**0.6237**	**0.9798**	**1.1193**

Table 2: Performance comparison of different variants in adversarial counterfactual distillation.

denote variants that remove adversarial learning and counterfactual distillation from forward distillation (FD) or reverse distillation (RD), respectively; variants *w/o AL* and *w/o CD* (i.e., $\gamma = 0$) remove adversarial learning and counterfactual distillation from both distillation directions, respectively; and *w/ CD*$_\gamma$ denotes using a fixed γ without updating it through adversarial learning. The MSE results are reported in Table 2. The following insights can be observed:

Compared to BiCoD, variants without adversarial learning (*w/o AL (FD)* and *w/o AL (RD)*) perform worse because they struggle to effectively learn bias suppression strength and overfit non-uniform rating distributions, thereby hindering effective bias suppression. Variants without counterfactual distillation (*w/o CD (FD)* and *w/o CD (RD)*) perform even worse, since they fail to address the contamination of true preference modeling caused by rating distribution bias during distillation. *w/o CD* performs the worst, highlighting that counterfactual distillation is essential for suppressing rating distribution bias. BiCoD outperforms *w/o CD*, validating the effectiveness of adversarial learning. BiCoD surpasses *w/ CD* γ, further demonstrating that adversarial learning can effectively adapt to the evolving student model, thereby facilitating counterfactual distillation on real-world datasets.

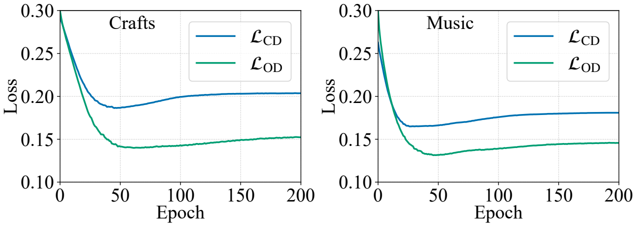


Figure 4: Loss trend comparison between $\mathcal{L}_{\text{CD}}$ and $\mathcal{L}_{\text{OD}}$.

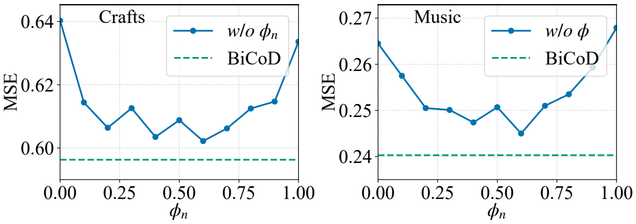


Figure 5: Sensitivity experiment on ϕ_n.

Loss Analysis of Counterfactual Distillation

Figure 4 illustrates the training dynamics of the adversarially optimized $\mathcal{L}_{\text{CD}}$ (mini-max) and the directly minimized $\mathcal{L}_{\text{OD}}$. In the early stage, θ_{MB} primarily focuses on feature construction and dominates the mini-max adversarial game, causing $\mathcal{L}_{\text{CD}}$ to decrease. In the later stages, $\mathcal{L}_{\text{CD}}$ remains higher than $\mathcal{L}_{\text{OD}}$ because adversarial learning suppresses the tendency of $\mathcal{L}_{\text{CD}}$ to exploit data biases to minimize loss, thereby increasing the difficulty of distillation and demonstrating the effectiveness of adversarial counterfactual distillation. The smooth trend of $\mathcal{L}_{\text{CD}}$ suggests that the adversarial mechanism does not compromise training stability, and its eventual convergence implies that a balance is achieved between online distillation and bias suppression.

Sensitivity Experiment

Figure 5 shows the model performance when scheduling coefficient ϕ_n is fixed (denoted as *w/o* ϕ_n) rather than scheduled by the cosine function. It can be observed that the model performs poorly when ϕ_n is set either too large or too small, as the weak distillation strength of RD/FD limits the ability to transfer knowledge or capture behavior-specific discrepancies. This confirms the effectiveness of the bidirectional distillation mechanism. Moreover, the model consistently underperforms with fixed values of ϕ_n compared to the dynamically scheduled BiCoD, validating the necessity of dynamically adjusting the distillation strength during training.

Conclusion

This paper proposes a bidirectional counterfactual distillation framework for review-based recommendation. It uses adversarial counterfactual distillation to prevent bias propagation through distillation from contaminating users' true preferences, and employs stage-aware bidirectional counterfactual distillation to avoid feature homogenization among behaviors. Extensive experiments on five real-world datasets demonstrate the effectiveness of the proposed framework.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant No. 62272143, U24A20332) and in part by the Seventh Special Support Plan for Innovation and Entrepreneurship in Anhui Province. The computation is completed on the HPC Platform of Hefei University of Technology.

References

Bauman, K.; Liu, B.; and Tuzhilin, A. 2017. Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 717–725. ACM.

Bengio, Y.; Louradour, J.; Collobert, R.; and Weston, J. 2009. Curriculum Learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, 41–48. Association for Computing Machinery.

Chen, C.; Zhang, M.; Liu, Y.; and Ma, S. 2018. Neural Attentional Rating Regression with Review-level Explanations. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, 1583–1592. ACM Press.

Ding, S.; Feng, F.; He, X.; Jin, J.; Wang, W.; Liao, Y.; and Zhang, Y. 2022. Interpolative Distillation for Unifying Biased and Debiased Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 40–49. ACM.

Du, H.; Yuan, H.; Zhao, P.; Zhuang, F.; Liu, G.; Zhao, L.; Liu, Y.; and Sheng, V. S. 2023. Ensemble Modeling with Contrastive Knowledge Distillation for Sequential Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 58–67. ACM.

Ganin, Y.; and Lempitsky, V. 2015. Unsupervised Domain Adaptation by Backpropagation. In Bach, F.; and Blei, D., eds., *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, 1180–1189. PMLR.

Glorot, X.; and Bengio, Y. 2010. Understanding the Difficulty of Training Deep Feedforward Neural Networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, 249–256. JMLR Workshop and Conference Proceedings.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative Adversarial Nets. In Ghahramani, Z.; Welling, M.; Cortes, C.; Lawrence, N.; and Weinberger, K., eds., *Advances in Neural Information Processing Systems*, volume 27, 1–9. Curran Associates, Inc.

Gou, J.; Chen, Y.; Yu, B.; Liu, J.; Du, L.; Wan, S.; and Yi, Z. 2024. Reciprocal Teacher-Student Learning via Forward and Feedback Knowledge Distillation. *IEEE Transactions on Multimedia*, 1–16.

Guo, Q.; Wang, X.; Wu, Y.; Yu, Z.; Liang, D.; Hu, X.; and Luo, P. 2020. Online Knowledge Distillation via Collaborative Learning. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11017–11026. IEEE.

He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 639–648. ACM.

Hou, Z.; Liu, X.; Cen, Y.; Dong, Y.; Yang, H.; Wang, C.; and Tang, J. 2022. GraphMAE: Self-Supervised Masked Graph Autoencoders. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '22, 594–604. Association for Computing Machinery.

Houyon, J.; Cioppa, A.; Ghunaim, Y.; Alfarra, M.; Halin, A.; Henry, M.; Ghanem, B.; and Van Droogenbroeck, M. 2023. Online Distillation With Continual Learning for Cyclic Domain Shifts. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2437–2446.

Ji, W.; Liu, X.; Zhang, A.; Wei, Y.; Ni, Y.; and Wang, X. 2023. Online Distillation-enhanced Multi-modal Transformer for Sequential Recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, 955–965. ACM.

Kingma, D. P.; and Ba, J. 2017. Adam: A Method for Stochastic Optimization. In *Proceedings of the 3rd International Conference for Learning Representations*, 1–15.

Kweon, W.; Kang, S.; and Yu, H. 2021. Bidirectional Distillation for Top-K Recommender System. In *Proceedings of the Web Conference 2021*, 3861–3871. ACM.

Lan, X.; Zhu, X.; and Gong, S. 2018. Knowledge Distillation by On-the-Fly Native Ensemble. In Bengio, S.; Wallach, H.; Larochelle, H.; Grauman, K.; Cesa-Bianchi, N.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 31, 1–11. Curran Associates, Inc.

Li, S.; Guo, D.; Liu, K.; Hong, R.; and Xue, F. 2023. Multimodal Counterfactual Learning Network for Multimedia-based Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1539–1548. ACM.

Li, Z.; Cheng, W.; Kshetramade, R.; Houser, J.; Chen, H.; and Wang, W. 2021. Recommend for a Reason: Unlocking the Power of Unsupervised Aspect-Sentiment Co-Extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 763–778. Association for Computational Linguistics.

Liu, D.; Cheng, P.; Lin, Z.; Luo, J.; Dong, Z.; He, X.; Pan, W.; and Ming, Z. 2022. KDCRec: Knowledge Distillation for Counterfactual Recommendation Via Uniform Data. *IEEE Transactions on Knowledge and Data Engineering*, 1–14.

Liu, D.; Li, J.; Du, B.; Chang, J.; and Gao, R. 2019. DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 344–352. ACM.

Liu, F.; Chen, H.; Cheng, Z.; Nie, L.; and Kankanhalli, M. 2023. Semantic-Guided Feature Distillation for Multimodal

Recommendation. In *Proceedings of the 31st ACM International Conference on Multimedia*, 6567–6575. ACM.

Liu, Y.; Yang, S.; Zhang, Y.; Miao, C.; Nie, Z.; and Zhang, J. 2021. Learning Hierarchical Review Graph Representations for Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 1–14.

Qiu, R.; Wang, S.; Chen, Z.; Yin, H.; and Huang, Z. 2021. CausalRec: Causal Inference for Visual Debiasing in Visually-Aware Recommendation. In *Proceedings of the 29th ACM International Conference on Multimedia*, 3844–3852. ACM.

Ren, Y.; Zhang, H.; Li, Q.; Fu, L.; Wang, X.; and Zhou, C. 2023. Self-Supervised Graph Disentangled Networks for Review-based Recommendation. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 2288–2295. International Joint Conferences on Artificial Intelligence Organization.

Shuai, J.; Zhang, K.; Wu, L.; Sun, P.; Hong, R.; Wang, M.; and Li, Y. 2022. A Review-aware Graph Contrastive Learning Framework for Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1283–1293.

Su, J.; Cao, J.; Liu, W.; and Ou, Y. 2021. Whitening Sentence Representations for Better Semantics and Faster Retrieval. *arXiv preprint arXiv:2103.15316*.

Wang, K.; Zhu, Y.; Zang, T.; Wang, C.; and Jing, M. 2024. Review-Enhanced Hierarchical Contrastive Learning for Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(8): 9107–9115.

Wang, K.; Zhu, Y.; Zang, T.; Wang, C.; Liu, K.; and Ma, P. 2023. Multi-Aspect Graph Contrastive Learning for Review-enhanced Recommendation. *ACM Transactions on Information Systems*, 3618106.

Wei, T.; Feng, F.; Chen, J.; Wu, Z.; Yi, J.; and He, X. 2021. Model-Agnostic Counterfactual Reasoning for Eliminating Popularity Bias in Recommender System. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 1791–1800. ACM.

Wu, G.; and Gong, S. 2021. Peer Collaborative Learning for Online Knowledge Distillation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(12): 10302–10310.

Wu, L.; He, X.; Wang, X.; Zhang, K.; and Wang, M. 2023. A Survey on Accuracy-oriented Neural Recommendation: From Collaborative Filtering to Information-rich Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 4425–4445.

Xia, L.; Huang, C.; and Zhang, C. 2022. Self-Supervised Hypergraph Transformer for Recommender Systems. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 10.

Xu, L.; Chia, Y. K.; and Bing, L. 2021. Learning Span-Level Interactions for Aspect Sentiment Triplet Extraction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 4755–4766. Association for Computational Linguistics.

Xv, G.; Liu, X.; Lin, C.; Li, H.; Li, C.; and Huang, Z. 2022. Lightweight Unbiased Multi-teacher Ensemble for Review-based Recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 4620–4624. ACM.

Yang, C.; An, Z.; Zhou, H.; Zhuang, F.; Xu, Y.; and Zhang, Q. 2023a. Online Knowledge Distillation via Mutual Contrastive Learning for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8): 10212–10227.

Yang, W.; Huo, T.; Liu, Z.; and Lu, C. 2023b. Review-Based Multi-intention Contrastive Learning for Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2339–2343. ACM.

Zhang, Y.; Xiang, T.; Hospedales, T. M.; and Lu, H. 2018. Deep Mutual Learning. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4320–4328. IEEE.

Zheng, L.; Noroozi, V.; and Yu, P. S. 2017. Joint Deep Modeling of Users and Items Using Reviews for Recommendation. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 425–434. ACM.

Zheng, Y.; Gao, C.; Li, X.; He, X.; Li, Y.; and Jin, D. 2021. Disentangling User Interest and Conformity for Recommendation with Causal Embedding. In *Proceedings of the Web Conference 2021*, 2980–2991. ACM.

Zhu, Z.; and Zhang, W. 2025. Exploring Feature-Based Knowledge Distillation for Recommender System: A Frequency Perspective. In Sun, Y.; Chierichetti, F.; Lauw, H. W.; Perlich, C.; Tok, W. H.; and Tomkins, A., eds., *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining, V.1, KDD 2025, Toronto, on, Canada, August 3-7, 2025*, 2182–2193. ACM.