Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Intent-Guided Diffusion with Adaptive Intent Fusion for User Cold-Start Recommendation

Lingxuan Li
Guang'an Institute of Technology
Luofei Jia
Guang'an Institute of Technology
Dafei Lin
Guang'an Institute of Technology
Chongmin Wang
Guang'an Institute of Technology
Yongxin Shi
shiyongxin123@nenu.edu.cn

Northeast Normal University

Research Article

Keywords: User cold-start recommendation, Diffusion model, Intent modeling, Adaptive fusion, Representation learning

Posted Date: June 29th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9277337/v1

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

Additional Declarations: No competing interests reported.

Intent-Guided Diffusion with Adaptive Intent Fusion for User Cold-Start Recommendation

Lingxuan Li[1], Luofei Jia[1], Dafei Lin[1], Chongmin Wang[1], Yongxin Shi[2]*
[1] College of Computer Engineering and Cyber Security, Guang'an Institute of Technology, Guang'an, China
[2] School of Information Science and Technology, Northeast Normal University, Changchun, China
*Corresponding author: Yongxin Shi
Email: shiyongxin123@nenu.edu.cn

Abstract

User cold-start recommendation aims to provide personalized recommendations for users with extremely limited interaction data, which remains a fundamental challenge in recommender systems. Existing methods mainly rely on auxiliary information, meta-learning, or sparsity simulation strategies. However, these approaches often depend heavily on limited observed behaviors and seldom exploit transferable latent intent patterns embedded in historical interaction data. To address this issue, we propose an Intent-Guided Diffusion Model (IGDM) for user cold-start recommendation. The proposed framework consists of three components: an intent extraction module, an adaptive intent fusion module, and an intent-guided diffusion module. Specifically, latent intent prototypes are first extracted from historical user interaction sequences to capture transferable preference patterns. Then, an adaptive intent fusion mechanism dynamically integrates the limited behavioral features of cold-start users with these intent prototypes according to their relevance, yielding a more reliable intent-aware representation. Finally, the fused intent representation is introduced as a conditional signal to guide the diffusion-based denoising process, allowing sparse and noisy user representations to be progressively refined toward more reliable preference estimation. Experiments on three real-world datasets, namely MovieLens, Last.FM, and Yelp, show that IGDM consistently outperforms multiple representative baselines on Recall, NDCG, and MAP, with more pronounced gains under sparse interaction settings. These results demonstrate the effectiveness of combining transferable intent modeling with diffusion-based representation refinement for user cold-start recommendation.

Keywords : User cold-start recommendation; Diffusion model; Intent modeling; Adaptive fusion; Representation learning

1 Introduction

The primary objective of recommender systems is to identify users' personalized preferences from a large pool of candidate items and recommend content that is likely to match their interests, thereby alleviating the problem of information overload [5,16,39]. In real-world applications, however,

platforms are frequently required to provide recommendations for users who have only a very limited number of historical interactions due to continuous user influx, business expansion, or short-term platform growth [32,36]. Such users are commonly referred to as cold-start users. Because their behavioral signals are sparse and incomplete, accurately modeling their preferences remains a challenging problem. As a result, user cold-start recommendation has become an important research topic in recommender systems [14,15,33].



Figure 1. An illustration of intents hidden in user interactions sequences.

To address this problem, early studies typically relied on auxiliary information, such as user profiles, item attributes, or side information collected during registration, to compensate for insufficient behavioral data [3,28]. However, such information is not always available in practical applications and may also be restricted by privacy constraints [29,44]. To reduce this dependency, some methods attempt to simulate sparse interaction scenarios and improve robustness under limited observations. For example, DropoutNet [17] applies dropout to user interaction embeddings to mimic missing behavioral evidence. Although effective to some extent, this strategy may further weaken already scarce collaborative signals and thus limit the model's ability to capture the initial preferences of cold-start users.

More recently, meta-learning-based methods have been introduced into user cold-start recommendation [1,40,41]. These methods learn transferable knowledge across multiple tasks and then rapidly adapt to new users with a small number of interactions. Representative examples include MeLU, MetaCS, MetaHIN, and TAPN [2,3,7,10]. Such methods have demonstrated stronger adaptability than traditional approaches. However, their performance still depends on the quality of limited support interactions. Under extremely sparse conditions, task-level adaptation may become unstable, and the learned initialization may not be sufficiently reliable for accurately characterizing new users, especially when their preferences differ from previously observed user patterns [12,30].

In many real-world recommendation scenarios, user behaviors such as clicks, browsing, and purchases are often driven by underlying intents [13,42]. For example, purchasing sports equipment may indicate a fitness-related intent, while browsing hotel or transportation information may reflect a travel-planning intent. These intents are usually implicit and embedded in interaction sequences

rather than directly observable. As illustrated in Figure 1, user interaction sequences with similar behavioral patterns can often be grouped together, and each group may correspond to a latent intent characterized by commonly co-occurring items. Motivated by this observation, recent studies have introduced intent modeling into recommendation systems. For instance, ICLRec [5] explores intent distributions within user behavior sequences and incorporates them into sequential recommendation. IOCRec [38] models user intent from both global and local perspectives to alleviate noise in contrastive learning, while ICSRec [6] further enhances representation learning through coarse-grained and fine-grained intent contrastive objectives. Although these methods have shown promising performance, they are primarily designed for scenarios with relatively sufficient interactions and do not directly address the severe sparsity issue in user cold-start recommendation.

Meanwhile, diffusion models [19,22,25–27] have demonstrated strong capacity for modeling complex data distributions in generative learning. Their core idea is to progressively perturb the original data with noise and then learn a reverse denoising process to recover informative representations or samples. Owing to this iterative refinement mechanism, diffusion models have recently been introduced into recommendation systems. Existing studies show that diffusion-based methods can effectively improve representation learning and capture complex behavioral patterns. However, most current diffusion-based recommendation approaches mainly focus on behavioral modeling itself and pay insufficient attention to the latent intent information embedded in historical interactions. As a result, their ability to recover reliable preference representations may still be limited under user cold-start settings, where observed behavioral evidence is extremely sparse and unstable.

From this perspective, two challenges remain insufficiently addressed in existing user cold-start recommendation methods. First, sparse user behaviors alone are often inadequate for learning stable and reliable preference representations. Second, although historical users contain transferable intent patterns, existing methods rarely leverage such intent priors in a way that can effectively guide representation refinement for cold-start users. Therefore, an effective cold-start recommendation model should not only exploit limited personalized behaviors, but also introduce transferable intent-level knowledge to compensate for sparse observations and progressively recover preference-consistent user representations.

To this end, we propose an Intent-Guided Diffusion Model (IGDM) for user cold-start recommendation. The proposed framework consists of three key components: an intent extraction module, an adaptive intent fusion module, and an intent-guided diffusion module. Specifically, we first extract latent intent prototypes from existing user interaction sequences to capture transferable preference patterns. Then, an adaptive intent fusion mechanism dynamically integrates the limited behavioral features of cold-start users with these intent prototypes according to their relevance, yielding a more reliable intent-aware representation. Finally, the fused intent representation is incorporated as a conditional signal to guide the diffusion-based denoising process, enabling sparse

and noisy user representations to be progressively refined toward their latent preferences. In this way, the proposed model jointly exploits transferable intent priors and iterative denoising refinement, making it particularly suitable for user cold-start recommendation under limited interaction data.

The main contributions of this work are summarized as follows:

- We propose an intent-guided diffusion framework for user cold-start recommendation. By integrating intent modeling with diffusion-based representation refinement, the proposed method improves preference learning under sparse interaction conditions.
- We design an adaptive intent fusion mechanism that dynamically weights latent intent prototypes according to their relevance to the cold-start user, thereby alleviating the instability of user representation learning caused by limited behavioral observations.
- We conduct extensive experiments on three real-world datasets. The results show that IGDM consistently achieves superior performance over multiple representative baselines, especially under highly sparse interaction settings, demonstrating its effectiveness for user cold-start recommendation.

2 RELATED WORK

This section reviews the studies most closely related to our work from three perspectives: user cold-start recommendation, intent-aware recommendation, and diffusion-based recommendation tasks.

2.1 User Cold-Start Recommendation

User cold-start recommendation aims to provide personalized recommendations for new users based on extremely limited historical interactions [1,4,21,42]. Since cold-start users lack sufficient behavioral data, accurately modeling their preferences under highly sparse conditions remains a fundamental challenge in recommender systems [43].

Existing studies on user cold-start recommendation can generally be categorized into two groups. The first group relies on auxiliary information to compensate for insufficient interaction data. For example, MeLU integrates user behaviors with side information to estimate user preferences for cold-start users [2,31], while related graph-based methods propagate preference signals through user–item interaction graphs in sparse recommendation scenarios [11]. These approaches can be effective when rich side information is available. However, in many real-world scenarios, auxiliary information is incomplete or unavailable due to data collection limitations and privacy constraints [29,44], which restricts the applicability of such methods.

The second group focuses on learning transferable knowledge from existing users. Meta-learning-based methods and pre-training strategies attempt to improve cold-start recommendation by transferring knowledge across users or tasks [20]. For instance, metaE constructs a task embedding space by learning similarities among tasks and enables fast adaptation to new users with only a few samples [3]. Similarly, DropoutNet simulates cold-start scenarios by randomly dropping part of the user interaction information during training, thereby improving robustness to sparse inputs [17]. Although these methods enhance adaptability under few-shot conditions, their performance still

heavily depends on the quality of limited observed interactions. Under extremely sparse settings, the learned representations may remain unstable and insufficient to characterize reliable user preferences.

In summary, existing cold-start recommendation methods either rely on external auxiliary information or depend on limited behavioral observations. Both strategies face inherent limitations under highly sparse conditions. In particular, most existing methods do not explicitly leverage transferable high-level semantic patterns embedded in historical user behaviors, especially latent intent structures, which could provide valuable prior knowledge for cold-start users.

2.2 Intent-Aware Recommendation

In recent years, increasing attention has been devoted to intent-aware recommendation due to its ability to capture deeper latent interest structures beyond explicit behavioral patterns [6,35,36,37]. Unlike traditional recommendation methods that mainly rely on observed interactions, intent-aware approaches aim to model the underlying motivations that drive user behaviors.

Several studies have explored intent modeling in recommendation systems. For example, ICLRec models intent distributions within user behavior sequences and incorporates them into a contrastive learning framework [5]. IOCRec further captures both global and local intent patterns to reduce noise interference and improve representation quality [39]. Related studies have also shown that integrating intent information can enhance the discriminative ability and robustness of recommendation models [30]. These methods demonstrate that latent intent information is beneficial for recommendation.

However, most existing intent-aware recommendation methods are developed for sequential recommendation scenarios where users have relatively sufficient interaction histories. Their effectiveness relies on the availability of sufficiently long behavior sequences to support stable intent extraction [5,38]. In user cold-start recommendation, by contrast, the target users have extremely limited interactions, making it difficult to directly infer reliable intent information from their own behaviors.

Therefore, directly applying existing intent-aware methods to cold-start scenarios is insufficient. A more suitable direction is to leverage intent patterns learned from existing users and transfer them to cold-start users as prior knowledge. How to effectively construct and utilize such transferable intent representations remains an open problem.

2.3 Diffusion Models for Recommendation

Diffusion models progressively corrupt original data with noise and then recover meaningful samples or representations through a reverse denoising process, demonstrating strong capability in modeling complex data distributions [8,22,23,27]. In recent years, diffusion models have achieved remarkable success in image generation, text generation, and related generative tasks [18,24], and have gradually been introduced into recommender systems.

In recommendation scenarios, existing studies have explored diffusion mechanisms for modeling user interest evolution and behavioral distributions. For example, Li et al. combined diffusion processes with Transformer architectures to model the dynamic evolution of user interests for sequential recommendation [29]. Wang et al. further proposed diffusion-based recommendation

models that improve recommendation efficiency and representation learning by incorporating temporal information into the diffusion process or by using encoders to reduce vector dimensionality [28]. These studies suggest that diffusion models possess strong representation optimization capability in recommendation tasks, especially when dealing with complex and noisy user behavior distributions.

Nevertheless, most existing diffusion-based recommendation methods mainly focus on behavioral patterns themselves and seldom explicitly model the latent intent information hidden in user interaction sequences. For user cold-start recommendation, relying solely on limited interaction behaviors for diffusion-based modeling may still be strongly affected by data sparsity. Therefore, integrating latent intent modeling with diffusion-based representation refinement has the potential to further improve the quality and stability of user representation learning in cold-start scenarios.

3 PRELIMINARIES

3.1 Problem Definition

Let U and V denote the sets of users and items, respectively. For any user $u \in U$ and item $v \in V$, the interaction sequence of user u can be constructed according to the chronological order of user-item interactions, which is denoted as $x_u = [x_u^1, x_u^2, \dots, x_u^n]$, where x_u^t denotes the item interacted by user u at the t-th time step, and n represents the length of the user's historical interaction sequence.

In user cold-start recommendation, cold-start users generally refer to users who have just entered the system or have only a few historical interactions. Following the common setting in previous work [7], we restrict the length of historical interaction sequences for cold-start users to no more than 20, and denote the interaction sequence of a cold-start user as x_c. Given a cold-start user u and its limited interaction sequence x_c, the goal of this paper is to learn the user's latent preference representation and further predict the next item that the user is most likely to interact with, thereby generating recommendation results that are consistent with the user's potential intent.

3.2 Diffusion Model

Diffusion Models (DMs) mainly consist of two parts: a forward noising process and a reverse denoising process. The core idea is to gradually inject noise into the original sample through the forward process so that it progressively approaches a simple distribution, and then recover a representation with the target distribution through the reverse denoising process.

3.2.1 Forward Noising Process

Let the original sample representation be denoted as s_0. In the forward diffusion process, the model gradually adds Gaussian noise to the sample through a Markov chain of length T. The noising process at step t is defined as

$$q(s_t \mid s_{t-1}) = \mathcal{N}\left(s_t; \sqrt{1-\beta_t}\, s_{t-1}, \beta_t I\right), \tag{1}$$

where $\beta_t \in (0,1)$ denotes the noise intensity at step t, and I is the identity matrix. As the diffusion steps increase, the sample representation is progressively perturbed by noise and eventually approaches an isotropic Gaussian distribution.

Based on the additive property of Gaussian distributions, s_t can also be directly expressed in terms of the original sample s_0 as

$$s_t = \sqrt{\bar{\alpha}_t}\, s_0 + \sqrt{1-\bar{\alpha}_t}\, \epsilon, \tag{2}$$

where $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$, and $\epsilon \sim \mathcal{N}(0, I)$.

3.2.2 Reverse Denoising Process

Corresponding to the forward process, the reverse process aims to progressively recover a representation that is closer to the original sample distribution from the noisy representation s_t. In practice, the denoising process is usually modeled by a parameterized neural network, whose conditional distribution can be written as

$$p_\theta(s_{t-1} \mid s_t) = \mathcal{N}(s_{t-1}; \mu_\theta(s_t, t), \Sigma_\theta(s_t, t)), \tag{3}$$

where $\mu_\theta(\cdot)$ and $\Sigma_\theta(\cdot)$ denote the mean and variance parameterized by θ, respectively. By progressively removing noise, the model is able to recover sample representations that are consistent with the target semantics from a simple noise distribution.

In practical training, diffusion models typically do not directly optimize the complete variational lower bound. Instead, they adopt a more stable noise prediction objective, in which the neural network predicts the noise ϵ injected during the forward process and minimizes the mean squared error between the ground-truth noise and the predicted noise:

$$\mathcal{L}_{diff} = \mathbb{E}_{t, s_0, \epsilon}[\| \epsilon - \epsilon_\theta(s_t, t) \|^2], \tag{4}$$

This objective effectively improves training stability and has therefore been widely adopted in diffusion-based models.

In this work, we further incorporate user intent information into the denoising process as a conditional signal, so that the generated user representations can better align with the latent preference patterns of cold-start users.

4 METHODOLOGY

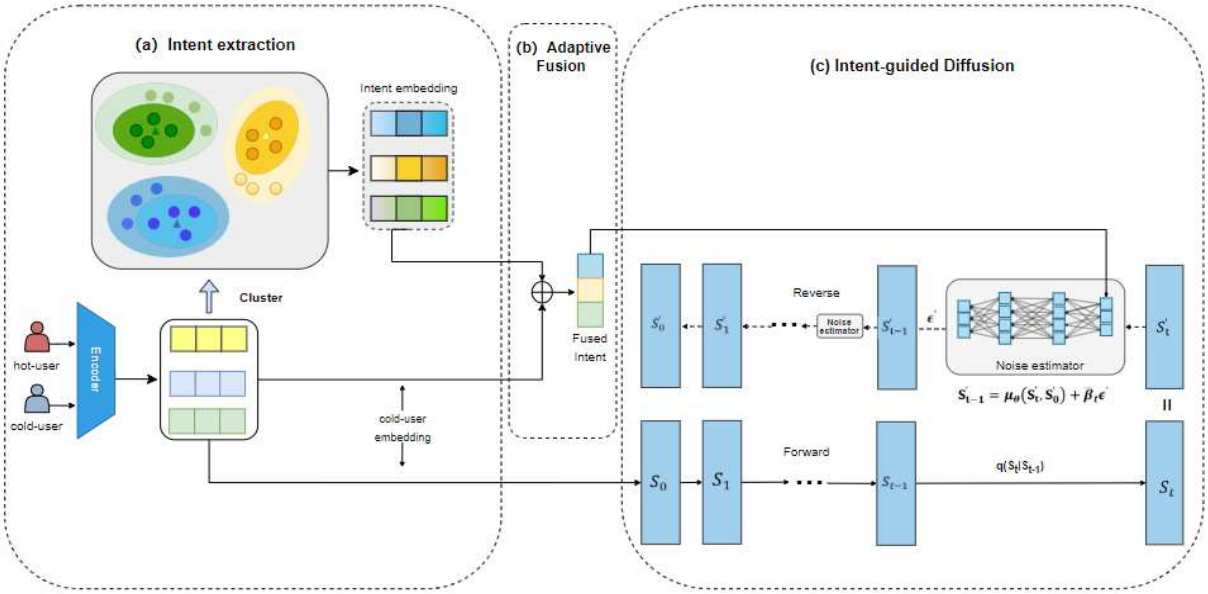


Figure 2. The framework of our proposed IGDM.

Figure 2 illustrates the overall framework of the proposed Intent-Guided Diffusion Model (IGDM). The model consists of three main components: the intent extraction module, the adaptive intent fusion module, and the intent-guided diffusion module.

The key idea of IGDM is to improve cold-start user representation learning by introducing transferable intent priors and refining sparse representations through a conditional diffusion process. Specifically, the intent extraction module first derives latent intent prototypes from historical user interaction sequences. Then, the adaptive intent fusion module dynamically integrates the limited behavioral features of cold-start users with these intent prototypes to construct a more reliable intermediate representation. Finally, the intent-guided diffusion module progressively denoises and refines the user representation under the guidance of the fused intent representation, producing a representation that better aligns with the user's latent preferences. The optimized representation is then used for next-item prediction.

4.1 Intent Extraction Module

User interaction behaviors are generally driven by latent intents, which are implicitly embedded in historical interaction sequences. Users with similar behavior patterns tend to share similar intent structures, while dissimilar behaviors usually correspond to different latent preference patterns. Based on this observation, we extract latent intent prototypes from existing user interaction sequences.

Let x_i and x_j denote the interaction sequences of users u_i and u_j, respectively. We first encode them into dense sequence representations S_i and S_j. The similarity between two sequences is measured using cosine similarit:

$$\text{Sim}\,(x_i, x_j) = \frac{S_i^\top S_j}{\| S_i \|_2 \| S_j \|_2}, \tag{5}$$

where S_i and S_j are the sequence representations of users u_i and u_j, respectively.

Based on the similarity measure, we apply K-means clustering to partition users into Kclusters, denoted as $C = \{C_1, C_2, \dots, C_K\}$. Each cluster represents a group of users with similar behavioral patterns and is regarded as a latent intent type. The intent prototype of cluster C_j is defined as:

$$I_j = \frac{1}{| C_j |} \sum_{i \in C_j} S_i, \tag{6}$$

where $| C_j |$ denotes the number of users in cluster C_j. The set of latent intent prototypes is denoted as $\mathcal{I} = \{I_1, I_2, \dots, I_K\}$.

These intent prototypes summarize transferable preference patterns from existing users and provide transferable intent priors for cold-start user modeling.

4.2 Adaptive intent fusion module

Cold-start users usually have extremely limited interaction data, making it difficult to obtain reliable preference representations from observed behaviors alone. To alleviate this issue, we design an

adaptive intent fusion module that incorporates transferable intent prototypes into user representation learning.

Let S_c denote the representation of a cold-start user u_c. To adaptively incorporate intent priors, we first compute the similarity between S_c and each intent prototype I_j, and then obtain normalized weights through a softmax function:

$$\omega_{cj} = \frac{\exp\left(\mathrm{Sim}\left(S_c, I_j\right)/\tau\right)}{\sum_{k=1}^{K} \exp\left(\mathrm{Sim}\left(S_c, I_k\right)/\tau\right)}, \tag{7}$$

where τ is a temperature parameter controlling the smoothness of the weight distribution.

The fused intent representation representation z_c is then defined as

$$z_c = \alpha S_c + (1-\alpha)\sum_{j=1}^{K} \omega_{cj}\, I_j, \tag{8}$$

where $\alpha \in [0,1]$ is a balancing coefficient controlling the trade-off between personalized behavioral features and intent priors. Although α is globally shared, the fusion process remains adaptive because the contribution of each intent prototype is dynamically determined according to its similarity to the cold-start user representation.

In this way, the model preserves limited personalized information while adaptively introducing intent-level prior knowledge. Different cold-start users can therefore receive different prototype combinations, which provides a more stable and informative conditional representation for the subsequent diffusion process.

4.3 Intent-Guided Diffusion Module

Although the fused representation already incorporates intent priors, it may still be unstable due to sparse observations and noisy behavioral signals. To further refine the representation, we design an intent-guided diffusion module based on the fused intent representation z_c. The module consists of a forward noising process and an intent-guided denoising process.

The motivation is that diffusion modeling performs iterative refinement rather than one-step reconstruction, which is more suitable for progressively recovering reliable preference representations from sparse and noisy inputs in cold-start scenarios.

4.3.1 Forward Noising

Let S_c^0 denote the initial representation of the cold-start user. The forward process progressively adds Gaussian noise through a Markov chain of length T:

$$q(S_c^t \mid S_c^{t-1}) = \mathcal{N}(S_c^t; \sqrt{1-\beta_t} S_c^{t-1}, \beta_t I), \tag{9}$$

where β_t controls the noise scale at step t.

Equivalently, S_c^t can be directly sampled from S_c^0 as:

$$S_c^t = \sqrt{\bar{\alpha}_t} S_c^0 + \sqrt{1-\bar{\alpha}_t}\epsilon, \tag{10}$$

where $\bar{\alpha}_t = \prod_{i=1}^{t}(1-\beta_i)$ and $\epsilon \sim \mathcal{N}(0, I)$.

4.3.2 Intent-Guided Denoising

In the reverse process, we incorporate the fused intent representation z_c as a conditional signal to guide denoising. A multilayer perceptron (MLP) is used as the noise estimator:

$$\hat{\epsilon}_c = \epsilon_\theta(S_c^t, t, z_c), \tag{11}$$

where $\epsilon_\theta(\cdot)$ denotes the learnable denoising network parameterized by θ.

To strengthen conditional guidance, we introduce a null intent vector $z_\emptyset$ and adopt classifier-free guidance[34]:

$$\hat{\epsilon}_c = (1+w)\epsilon_\theta(S_c^t, t, z_c) - w\epsilon_\theta(S_c^t, t, z_\emptyset), \tag{12}$$

where w controls the strength of intent guidance.

The mean of the reverse process is defined as:

$$\mu_\theta(S_c^t, t) = \frac{1}{\sqrt{1-\beta_t}}\left(S_c^t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\epsilon}_c\right), \tag{13}$$

and the denoising step is:

$$S_c^{t-1} = \mu_\theta(S_c^t, t) + \sigma_t\eta, \eta \sim \mathcal{N}(0, I), \tag{14}$$

where $\sigma_t$is the variance term at step t.

Through iterative denoising, the model progressively refines the noisy representation, enabling it to recover preference signals that are more consistent with the user's latent interests.

4.4 Loss Function

We optimize the diffusion model by minimizing the noise prediction error:

$$\mathcal{L}_{diff} = \mathbb{E}_{t,S_c^0,\epsilon}[\| \epsilon - \epsilon_\theta(S_c^t, t, z_c) \|^2], \tag{15}$$

After obtaining the final user representation h_c, the matching score between user u_c and item v is computed as:

$$\hat{y}_{cv} = h_c^\top e_v, \tag{16}$$

where e_v is the embedding of item v.

We adopt the Bayesian Personalized Ranking (BPR) loss:

$$\mathcal{L}_{rec} = -\sum \log \sigma(\hat{y}_{cv^+} - \hat{y}_{cv^-}), \tag{17}$$

where v^+ and v^-denote positive and negative items, respectively.

The final objective function is defined as:

$$\mathcal{L} = \mathcal{L}_{rec} + \lambda\mathcal{L}_{diff}, \tag{18}$$

where λ balances recommendation learning and diffusion learning.

5 EXPERIMENTS

In this section, we aim to answer the following research questions:

- RQ1: How does IGDM perform compared with existing user cold-start recommendation methods?
- RQ2: How do different key modules contribute to the overall performance of the model?
- RQ3: How do the key hyperparameters affect model performance?
- RQ4: Can IGDM maintain stable advantages under different cold-start levels?

- RQ5: Beyond recommendation accuracy, can IGDM learn more discriminative user representations?

5.1 Experimental Setup

5.1.1 Datasets

To comprehensively evaluate the effectiveness of the proposed model, we conduct experiments on three public datasets: MovieLens-1M, Last.FM, and Yelp.

- MovieLens-1M: This dataset contains users' movie rating records and provides auxiliary information such as user attributes and item attributes. It is a widely used benchmark dataset in recommender system research.
- Last.FM: This dataset records users' music listening behaviors on the Last.FM platform and is commonly used to evaluate user interest modeling and sequential recommendation methods.
- Yelp: This dataset contains user reviews and interaction behaviors with merchants, covering multiple types of information such as business profiles, user information, and review contents, and thus has strong practical relevance.

The statistics of the three datasets are summarized in Table 1.

Table 1. Datasets.

Dataset	#User	#Item	#Interactions
MovieLens	6040	3529	1000209
Last.FM	1872	3846	42346
Yelp	5883	4712	216584

5.1.2 Data Split

We adopt a time-aware data splitting strategy to simulate realistic recommendation scenarios. For each user, interactions are first sorted chronologically and then divided into training, validation, and test sets with a ratio of 8:1:1.

To construct the user cold-start setting, we first identify users with limited interaction histories. Specifically, users whose total number of interactions does not exceed 20 are treated as cold-start users. For these users, only a small subset of their earliest interactions is used as observed behavior, while the remaining subsequent interactions are used for evaluation.

Importantly, latent intent prototypes are learned only from training users, and cold-start users are excluded from the prototype extraction process. This design avoids information leakage and ensures that the learned intent priors are transferable rather than directly derived from target users.

All evaluations are conducted strictly on interactions that occur after the observed behaviors, thereby ensuring a fair and realistic cold-start evaluation protocol.

5.1.3 Evaluation Metrics.

User cold-start recommendation can be regarded as a Top-N recommendation task. We adopt three widely used ranking metrics for evaluation: Recall@ k, NDCG@ k, and MAP@ k.

- **Recall@** k measures the proportion of ground-truth interacted items that appear in the Top-N recommendation list, reflecting the model's ability to cover user interests.

- **NDCG@k** considers both the relevance of recommended items and their ranked positions, where a higher value indicates that the ranking order is more consistent with user preferences.
- **MAP@k** computes the average precision of the recommendation list, reflecting the overall ranking quality of the model.

In this work, we report Recall@k, NDCG@k, and MAP@k under k = 5 and 10.

5.1.4 Baselines

We compare IGDM with three categories of representative baselines, including traditional cold-start recommendation methods, meta-learning-based cold-start recommendation methods, and sequential recommendation methods.

Traditional cold-start recommendation methods

- **GC-MC [45]:** GC-MC learns user and item embeddings through graph convolution operations to capture complex user-item relationships.
- **DropoutNet [17]:** DropoutNet simulates cold-start scenarios by randomly dropping user interaction information, thereby improving robustness under sparse settings.

Meta-learning-based cold-start recommendation methods

- **MeLU [2]:** MeLU first learns transferable initialization parameters through meta-learning and then quickly adapts to new users with limited interaction data.
- **MetaCS [3]:** MetaCS improves cold-start recommendation by learning task-level transferable knowledge in a meta-learning framework.
- **MetaHIN [10]:** MetaHIN leverages rich semantic information in heterogeneous information networks to provide more accurate prior knowledge for new users.
- **TAPN [7]:** TAPN learns globally shared initialization parameters and then performs local fine-tuning for cold-start users.
- **ColdNas [9]:** ColdNas employs neural architecture search to dynamically configure user-specific model structures for prediction.

Sequential recommendation methods

- **SelfGNN [46]:** SelfGNN captures high-order collaborative signals in short-term user behaviors through graph neural networks and combines them with sequential modeling and contrastive learning to enhance representation robustness.
- **DCRec [47]:** DCRec employs a multi-channel network to model users' conformity behaviors in sequential interactions and dynamically strengthens contrastive learning between sequence encoding and graph-based user interest modeling.

Although recent diffusion-based recommendation models have shown promising results, most of them are designed for dense interaction scenarios and do not explicitly target user cold-start settings. Therefore, we focus on comparing with representative cold-start and sequential baselines.

5.2 Overall Comparison (RQ1)

To validate the effectiveness of the proposed IGDM, we compare it with multiple representative baselines on three public datasets, namely MovieLens-1M, Last.FM, and Yelp. The results are

reported in Table 2, where the best results are highlighted in bold, the second-best results are underlined, and Improve (%) denotes the relative improvement of IGDM over the best baseline.

Overall, IGDM achieves consistent but moderate improvements over strong baselines across all three datasets. The gains are more pronounced under sparse interaction conditions, which indicates that the proposed model is effective in alleviating performance degradation caused by insufficient behavioral data in cold-start recommendation.

From the perspective of different categories of baselines, traditional methods such as GC-MC and DropoutNet exhibit relatively weak performance. This is mainly because they rely heavily on sufficient user-item interaction signals, which are difficult to obtain in cold-start scenarios.

Sequential recommendation methods, including SelfGNN and DCRec, achieve better results by exploiting structural and sequential information. However, when user interaction sequences are extremely short, their ability to capture stable behavioral patterns is still limited.

Meta-learning-based methods, including MeLU, MetaCS, MetaHIN, TAPN, and ColdNas, generally show stronger performance than traditional methods, since they can learn transferable knowledge across tasks. Nevertheless, under extremely sparse conditions, these methods may still suffer from unstable adaptation or insufficient support information.

Compared with these baselines, IGDM introduces transferable intent prototypes and further refines sparse user representations through an intent-guided diffusion process. This design enables the model to better exploit global behavioral patterns while suppressing the noise introduced by sparse interactions.

Compared with meta-learning approaches, IGDM avoids unstable task-level adaptation and instead relies on more stable intent priors. Compared with sequential models, it reduces the dependency on sequence length. Compared with traditional graph-based methods, it explicitly models latent preference structures and progressively refines them through denoising.

Overall, the results demonstrate that integrating latent intent modeling with diffusion-based representation refinement is an effective solution for user cold-start recommendation, especially in sparse data environments.

Table 2. Overall performance comparison between baselines and IGDM on three datasets. The best results are highlighted in bold, while the second-best are underlined. Percent improvements (%), indicating the relative enhancement of IGDM over the best baseline result, are also provided.

	N@5	N@10	R@5	R@10	M@5	M@10	N@5	N@10	R@5	R@10	M@5	M@10	N@5	N@10	R@5	R@10	M@5	M@10
	MovieLens-1M						Last.fm						Yelp					
GC-MC	0.1083	0.1097	0.0223	0.0341	0.0974	0.0973	0.0875	0.0903	0.0107	0.0206	0.0684	0.0798	0.0905	0.1026	0.0153	0.0223	0.0767	0.0687
DropoutNet	0.1173	0.1167	0.0265	0.0407	0.1058	0.1058	0.0881	0.0925	0.0113	0.0251	0.0759	0.0908	0.0921	0.1186	0.0168	0.0249	0.0813	0.0731
DCRec	0.1205	0.1254	0.0273	0.0411	0.1112	0.1216	0.0893	0.0929	0.0119	0.0256	0.0764	0.0917	0.0933	0.1201	0.0175	0.0257	0.0847	0.0746
MetaHIN	0.1216	0.1281	0.0283	0.0423	0.1116	0.1219	0.0901	0.0936	0.0125	0.0269	0.0803	0.0953	0.1056	0.1293	0.0193	0.0261	0.0853	0.0753
MeLU	0.1223	0.1285	0.0288	0.0429	0.1121	0.1233	0.0907	0.0987	0.0129	0.0265	0.0796	0.0987	0.1058	0.1306	0.0209	0.0277	0.0982	0.1028
TAPN	0.1228	0.1291	0.0291	0.0435	0.1127	0.1237	0.0912	0.1011	0.0137	0.0281	0.0818	0.1035	0.1067	0.1349	0.0212	0.0283	0.103	0.1073
SelfGNN	0.1231	0.1293	0.0295	0.0446	0.1135	0.1243	0.0918	0.0995	0.0132	0.0276	0.0826	0.1046	0.1075	0.1352	0.0221	0.0294	0.1015	0.1086
ColdNas	0.1237	0.1305	0.0298	0.0441	0.1148	0.1248	0.0927	0.1005	0.0148	0.0287	0.0812	0.1048	0.1132	0.1381	0.0225	0.0318	0.1023	0.1176
IGDM	**0.1263**	**0.1356**	**0.0313**	**0.0462**	**0.1167**	**0.1289**	**0.0954**	**0.1047**	**0.0156**	**0.0301**	**0.0843**	**0.1094**	**0.1151**	**0.1436**	**0.0232**	**0.0324**	**0.1064**	**0.1213**
Improve(%)	2.10%	3.91%	5.03%	3.59%	1.66%	3.28%	2.91%	3.56%	5.41%	4.88%	2.06%	4.38%	1.68%	3.98%	3.11%	1.89%	4.01%	3.15%

5.3 Ablation Study (RQ2)

To verify the contribution of each key module to the overall performance, we construct two variant models and compare them with the full IGDM under the same experimental setting:

- w/o IF (intention-fusion): removing the intent fusion module while retaining only the diffusion-based representation learning process;
- w/o IG (intent-guided): removing the intent-guided mechanism so that the diffusion denoising process is no longer constrained by the fused intent representation.

The ablation results are shown in Figure 3.

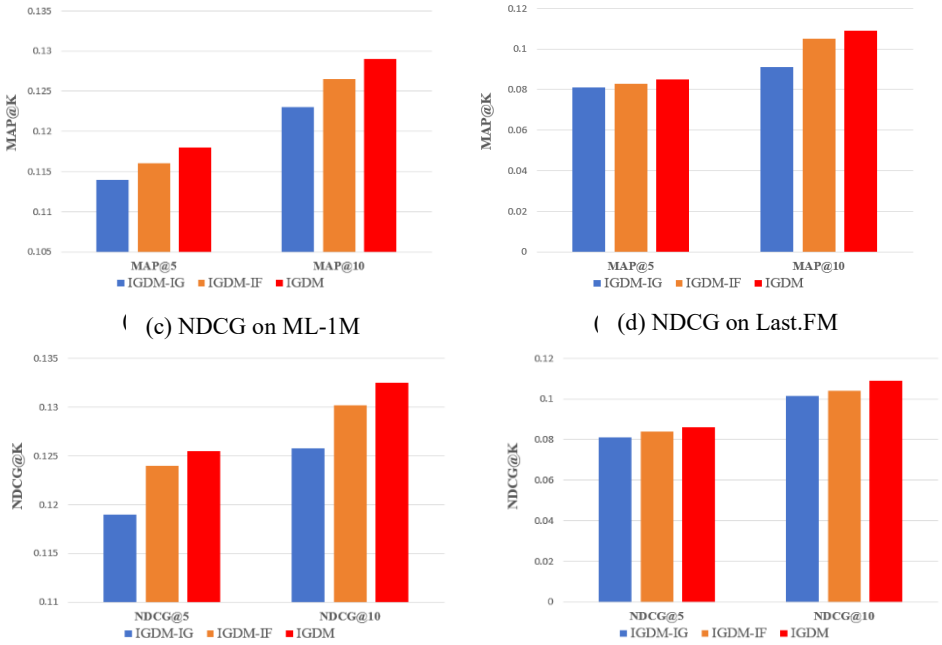


Figure 3. Ablation experiments.

It can be observed that removing either component leads to clear performance degradation, indicating that both modules contribute positively to the final recommendation results.

In particular, removing the intent-guided mechanism causes a larger performance drop. This suggests that explicit intent guidance plays an important role in stabilizing the diffusion-based denoising process under sparse inputs and helps the model recover preference representations that are more consistent with users' underlying interests.

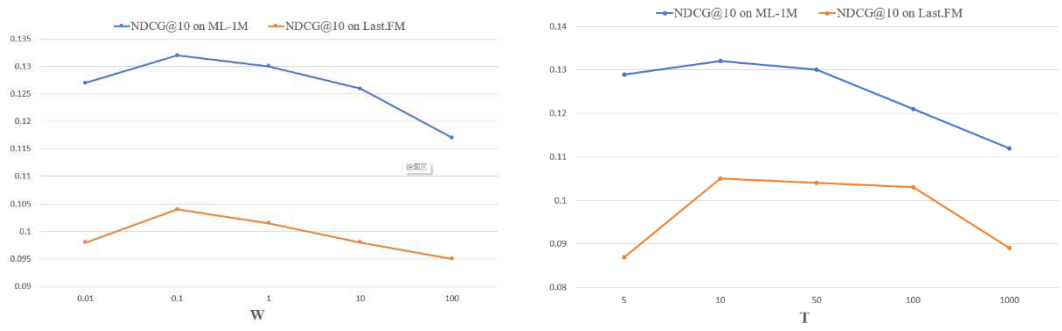
In addition, removing the intent fusion module also leads to noticeable deterioration. This indicates that the limited behavioral information of cold-start users is insufficient to directly support high-quality representation learning. By incorporating transferable intent priors extracted from existing users, the model can alleviate the instability caused by sparse interactions.

Although we do not explicitly include a variant without the diffusion process, such a comparison would provide a more direct assessment of the necessity of diffusion-based refinement. Overall, the

ablation study verifies that both the intent fusion module and the intent-guided diffusion mechanism are beneficial to the proposed framework.

5.4 Hyper-parameter Analysis (RQ3)

To further analyze the sensitivity of the proposed model to key hyperparameters, we investigate the effects of the diffusion step number T and the intent guidance weight w on model performance. The experiments are conducted on the MovieLens-1M dataset.



(a) W (b) T

Figure 4. Hyper-parameter Analysis.

First, with respect to the diffusion step number T, the results show that model performance first improves and then declines as T increases, reaching the best performance at around T = 50. This indicates that an appropriate number of diffusion steps helps the model better capture the latent structure of the data. However, when the diffusion process becomes too long, the original user behavioral signal is excessively diluted, which increases the difficulty of recovering useful information during denoising and thus leads to performance degradation. In contrast, when T is too small, the injected perturbation is insufficient and the model cannot adequately explore the latent representation space.

Second, regarding the intent guidance weight w, the results show that model performance steadily improves as w increases and reaches the optimum at w = 0.1. When w is too small, the model cannot fully exploit the fused intent representation to constrain the denoising process. On the other hand, if w is too large, the conditional intent may exert overly strong influence on the generation process, thereby weakening the preservation of the original behavioral patterns and impairing the quality of the learned representation.

Overall, under the current experimental setting, the best hyperparameter configuration is achieved when the diffusion step number is set to T = 50 and the intent guidance weight is set to w = 0.1. This configuration allows the model to better leverage intent guidance while preserving essential behavioral characteristics, thereby leading to superior recommendation performance.

5.5 Cold-start Sensitivity Analysis (RQ4)

To further evaluate the recommendation performance of the proposed model under different levels of cold-start severity, we conduct a cold-start sensitivity analysis on the MovieLens-1M dataset. Specifically, test users are divided into four levels according to the number of their historical

interactions, as shown in Table 3. Among them, Level 1 represents severely cold-start users, while Level 4 is closer to a less sparse recommendation setting. This partition provides a finer-grained view of model behavior under different levels of data sparsity.

Table 3. Cold-start level partition.

Level	Interaction range	#Test samples
Level 1	[1-5)	300
Level 2	[5-10)	300
Level 3	[10-20)	350
Level 4	[20-50)	350

For comparison, we select three competitive baselines, namely DropoutNet, SelfGNN, and ColdNas. The results are reported in Table 4.

From the overall trend, as the cold-start level changes from Level 1 to Level 4, all models show increasing performance in terms of Recall@10, NDCG@10, and MAP@10. This indicates that as user interaction information becomes richer, all methods are able to learn more sufficient preference patterns and thus achieve better recommendation results.

More importantly, IGDM achieves the most significant improvements under the highly sparse settings of Level 1 and Level 2, where user interaction signals are extremely limited. This indicates that the proposed model is particularly effective when observed behavior is insufficient to support stable preference estimation.As the number of user interactions increases, the performance gap between IGDM and strong baselines gradually narrows. This suggests that the advantage of intent-guided diffusion is most prominent in severe cold-start scenarios, while its relative gain becomes smaller when richer behavioral information is available.

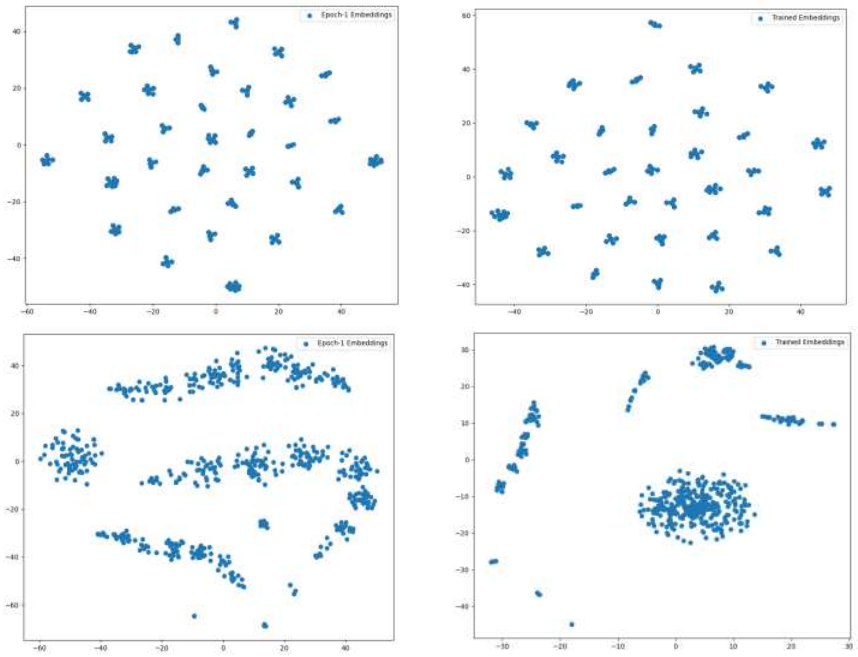
These results confirm that IGDM is especially suitable for data-scarce cold-start settings, while still maintaining competitive performance in relatively less sparse conditions.

Table 4. Cold-start sensitivity analysis results.

Cold-start level	Model	Recall@10	NDCG@10	MAP@10
Level 1	DropoutNet	0.0203	0.0816	0.0907
	SelfGNN	0.0221	0.0847	0.0928
	ColdNas	0.0257	0.0875	0.0955
	IGDM	0.0269	0.0891	0.0965
Level 2	DropoutNet	0.0429	0.1395	0.1267
	SelfGNN	0.0448	0.1411	0.1283
	ColdNas	0.0457	0.1418	0.1308
	IGDM	0.0475	0.1427	0.1327
Level 3	DropoutNet	0.0463	0.1439	0.1616
	SelfGNN	0.0501	0.1611	0.1751
	ColdNas	0.0498	0.1569	0.1695
	IGDM	0.0513	0.1637	0.1789
Level 4	DropoutNet	0.0634	0.1957	0.1579
	SelfGNN	0.0835	0.2534	0.2138
	ColdNas	0.0731	0.2106	0.1963
	IGDM	0.0849	0.2537	0.2154

5.6 Visual Analysis (RQ5)

To further analyze the effectiveness of IGDM in user representation learning, we employ t-SNE to visualize the high-dimensional user representations learned by different models. Specifically, we compare the representation distributions of ColdNas and IGDM on the Last.FM dataset, and the results are shown in Figure 5.



(a)Epoch-1 Embeddings (b) Trained Embeddings

Figure 5. User embeddings visualized by t-SNE. The top row shows the results of ColdNas, while the bottom row shows the results of IGDM.

More specifically, we randomly sample 200 user representation vectors from each model and project them into a two-dimensional space for visualization. By comparing the resulting distributions, it can be observed that the user representations learned by ColdNas are relatively scattered and mainly form several local clusters. In contrast, the user representations learned by IGDM exhibit a clearer structural distribution, forming a more compact main cluster while still preserving a certain degree of boundary separation for users with distinct personalized preferences.

This phenomenon suggests that, under the guidance of the intent-guided diffusion mechanism, IGDM is able to capture latent similarities among users more effectively while preserving individual differences, thereby producing more stable and discriminative user representations.

It should be noted that t-SNE is mainly used here for qualitative analysis, and the visualization results should be interpreted with caution. Overall, the visual results still provide supportive evidence that IGDM learns representations with a better balance between clustering structure and personalization, which helps explain its performance gains in recommendation tasks.

6 CONCLUSION

In this paper, we investigate the problem of user cold-start recommendation, where user interaction data is extremely limited and preference modeling is therefore highly challenging. To address this issue, we propose an Intent-Guided Diffusion Model (IGDM), which integrates transferable intent modeling with diffusion-based representation refinement for cold-start user recommendation.

Specifically, the proposed framework first extracts latent intent prototypes from existing user interaction sequences to capture transferable preference patterns. Then, an adaptive intent fusion module is introduced to dynamically incorporate these intent prototypes into the representation learning process of cold-start users according to their relevance, thereby alleviating the instability caused by sparse behavioral observations. Based on the fused intent representation-aware representation, an intent-guided diffusion mechanism is further employed to progressively refine noisy and incomplete user representations through conditional denoising, enabling them to better reflect users' latent preferences.

Extensive experiments on three real-world datasets demonstrate that IGDM consistently outperforms multiple representative baselines in terms of Recall, NDCG, and MAP. The improvements are especially evident under highly sparse interaction settings, indicating that the proposed method is particularly effective for severe cold-start scenarios. In addition, ablation studies confirm that both the intent fusion module and the intent-guided diffusion module contribute positively to the final recommendation performance.

Although the proposed method achieves promising results, several limitations remain. First, the current framework focuses only on user cold-start recommendation and has not yet been extended to item cold-start or more general dual cold-start settings. Second, the intent prototypes are constructed from historical user behaviors, and their transferability may still be limited when new users exhibit preference patterns that differ significantly from previously observed groups. Third, the diffusion-based refinement process introduces additional computational overhead, which may affect scalability in large-scale recommendation scenarios.

In future work, we will investigate more efficient intent modeling and diffusion mechanisms to reduce computational cost, and further extend the proposed framework to broader cold-start recommendation settings.

Declarations

Funding

This research received no external funding.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

Lingxuan Li: Conceptualization, Methodology, Investigation, Experiments, Writing – Original Draft.

Yongxin Shi: Supervision, Writing – Review & Editing.

Luofei Jia: Data Curation, Validation.

Dafei Lin: Data Curation, Validation.

Chongmin Wang: Data Curation, Validation.

Ethics Approval

This study does not involve human participants or animals. Ethical approval is not required.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Data Availability

The datasets used in this study are publicly available, including MovieLens-1M, Last.FM, and Yelp. These datasets can be obtained from their corresponding public sources.

References

[1] Bharadhwaj H. Meta-learning for user cold-start recommendation[C]//2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019: 1-8.

[2] Lee H, Im J, Jang S, et al. MeLU: Meta-learned user preference estimator for cold-start recommendation[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019: 1073-1082.

[3] Zhu Y, Xie R, Zhuang F, et al. Learning to warm up cold item embeddings for cold-start recommendation with meta scaling and shifting networks[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021: 1167-1176.

[4] Salha-Galvan G, Hennequin R, Chapus B, et al. Cold start similar artists ranking with gravity-inspired graph autoencoders[C]//Proceedings of the 15th ACM Conference on Recommender Systems. 2021: 443-452.

[5] Yongjun Chen, Zhiwei Liu, Jia Li, Julian J. McAuley, and Caiming Xiong. 2022.Intent Contrastive Learning for Sequential Recommendation. In WWW. 2172–2182.

[6] Qin X, Yuan H, Zhao P, et al. Intent Contrastive Learning with Cross Subsequences for Sequential Recommendation[C]//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. 2024: 548-556.

[7] Xixun Lin, Jia Wu, Chuan Zhou, Shirui Pan, Yanan Cao, and Bin Wang. 2021.Task-adaptive neural process for user cold-start recommendation. In The Web Conference. 1306–1316.

[8] Lin J, Cao Y, Yu Y, et al. Diffusion Models for Recommender Systems: From Content Distribution To Content Creation[C]//Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2. 2025: 6074-6085.

[9] Wu S, Wang Y, Jing Q, et al. ColdNAS: Search to Modulate for User Cold-Start Recommendation[C]//Proceedings of the ACM Web Conference 2023. 2023: 1021-1031.

[10] Lu Y, Fang Y, Shi C. Meta-learning on heterogeneous information networks for cold-start recommendation[C]//Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining. 2020: 1563-1573.

[11] Wang Y, Han L, Qian Q, et al. Personalized recommendation via multi-dimensional meta-paths temporal graph probabilistic spreading[J]. Information Processing & Management, 2022, 59(1): 102787.

[12] Chen H, Wang Z, Huang F, et al. Generative adversarial framework for cold-start item recommendation[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022: 2565-2571.

[13] Li H, Wang Y, Xiao Z, et al. DisCo: graph-based disentangled contrastive learning for cold-start cross-domain recommendation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2025, 39(11): 12049-12057.

[14] Lai P L, Chen C Y, Lo L W, et al. ColdGAN: Resolving Cold Start User Recommendation by using Generative Adversarial Networks[J]. arXiv preprint arXiv:2011.12566, 2020.

[15] Zheng J, Ma Q, Gu H, et al. Multi-view denoising graph auto-encoders on heterogeneous information networks for cold-start recommendation[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021: 2338-2348.

[16] Changfeng Sun, Han Liu, Meng Liu, Zhaochun Ren, Tian Gan, and Liqiang Nie.2020. LARA: Attribute-to-feature adversarial learning for new-item recommendation. In WSDM.

[17] Volkovs M, Yu G, Poutanen T. DropoutNet: Addressing cold start in recommender systems[J]. Advances in neural information processing systems, 2017, 30.

[18] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition. 2022: 10684-10695.

[19] Bui T N, Nguyen H S, Nguyen C V T, et al. Personalized diffusion model reshapes cold-start bundle recommendation[C]//Companion Proceedings of the ACM on Web Conference 2025. 2025: 3088-3091.

[20] Hao B, Yin H, Zhang J, et al. A Multi-strategy-based Pre-training Method for Cold-start Recommendation[J]. ACM Transactions on Information Systems, 2023, 41(2): 1-24.

[21] Du Y, Zhu X, Chen L, et al. MetaKG: Meta-learning on knowledge graph for cold-start recommendation[J]. IEEE Transactions on knowledge and data engineering, 2022, 35(10): 9850-9863.

[22] Wei Y, Wang X, Li Q, et al. Contrastive learning for cold-start recommendation[C]//Proceedings of the 29th ACM International Conference on Multimedia. 2021: 5382-5390.

[23] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising diffusion probabilistic models. In Advances in Neural Information Processing Systems,Vol. 33. 6840–6851.

[24] Dhariwal P, Nichol A. Diffusion models beat gans on image synthesis[J]. Advances in neural information processing systems, 2021, 34: 8780-8794.

[25] Harper F M, Konstan J A. The movielens datasets: History and context[J]. ACM transactions on interactive intelligent systems (tiis), 2015, 5(4): 1-19.

[26] Li X, Thickstun J, Gulrajani I, et al. Diffusion-lm improves controllable text generation[J]. Advances in neural information processing systems, 2022, 35: 4328-4343.

[27] Smith K E, Smith A O. Conditional GAN for timeseries generation[J]. arXiv preprint arXiv:2006.16477, 2020.

[28] Zou H, Kim Z M, Kang D. Diffusion models in nlp: A survey[J]. arXiv preprint arXiv:2305.14671, 2023.

[29] Wang W, Xu Y, Feng F, et al. Diffusion recommender model[C]//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2023: 832-841.

[30] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Li, Jay Whang, Emily L Denton, Kamyar Ghasemipour, Raphael Gontijo Lopes, Burcu Karagol Ayan, Tim Salimans, et al. 2022. Photorealistic text-to-image diffusion models with deep language understanding. Advances in Neural Information Processing Systems 35(2022), 36479–36494.

[31] Hu J, An T, Yu Z, et al. Contrastive learning for cold start recommendation with adaptive feature fusion[C]//2025 5th International Conference on Consumer Electronics and Computer Engineering (ICCECE). IEEE, 2025: 520-524.

[32] Debashis Das, Laxman Sahoo, and Sujoy Datta. 2017. A survey on recommendation system. International Journal of Computer Applications 160, 7 (2017).

[33] Jyotirmoy Gope and Sanjay Kumar Jain. 2017. A survey on solving cold start problem in recommender systems. In 2017 International Conference on Computing,Communication and Automation (ICCCA). IEEE, 133–138.

[34] Jonathan Ho and Tim Salimans. 2022. Classifier-free diffusion guidance. arXiv preprint arXiv:2207.12598 (2022).

[35] Chang B, Karatzoglou A, Wang Y, et al. Latent user intent modeling for sequential recommenders[C]//Companion Proceedings of the ACM Web Conference 2023. 2023: 427-431.

[36] Renqin Cai, Jibang Wu, Aidan San, Chong Wang, and Hongning Wang. 2021. Category-aware Collaborative Sequential Recommendation. In SIGIR. 388–397.

[37] Wanyu Chen, Pengjie Ren, Fei Cai, Fei Sun, and Maarten de Rijke. 2020. Improving End-to-End Sequential Recommendations with Intent-aware Diversification. In CIKM. 175–184.

[38] Shoujin Wang, Liang Hu, Yan Wang, Quan Z Sheng, Mehmet Orgun, and Long-bing Cao. 2019. Modeling multi-purpose sessions for next-item recommendations via mixture-channel purpose routing networks. In IJCAI. 3771–3777.

[39] Xuewei Li, Aitong Sun, Mankun Zhao, Jian Yu, Kun Zhu, Di Jin, Mei Yu, and Ruiguo Yu. 2023. Multi-Intention Oriented Contrastive Learning for Sequential Recommendation. In WSDM. 411–419.

[40] Pigi Kouki, Shobeir Fakhraei, James Foulds, Magdalini Eirinaki, and Lise Getoor. 2015. Hyper: A flexible and extensible probabilistic framework for hybrid recommender systems. In ACM Conference on Recommender Systems (RecSys).

[41] Yaqing Wang, Quanming Yao, James T Kwok, and Lionel M Ni. 2020. Generalizing from a few examples: A survey on few-shot learning. Comput. Surveys 53, 3(2020), 1–34.

[42] Requena B, Cassani G, Tagliabue J, et al. Shopper intent prediction from clickstream e-commerce data with minimal browsing information[J]. Scientific reports, 2020, 10(1): 16983.

[43] Mattia Bianchi, Federico Cesaro, Filippo Ciceri, Mattia Dagrada, Alberto Gasparin, Daniele Grattarola, Ilyas Inajjar, Alberto Maria Metelli, and Leonardo Cella. 2017. Content-based approaches for cold-start job recommendations. In Proceedings of the Recommender Systems Challenge.

[44] Yu Xin and Tommi Jaakkola. 2014. Controlling privacy in recommender systems. In Annual Conference on Neural Information Processing Systems (NeurIPS).

[45] Berg R, Kipf T N, Welling M. Graph convolutional matrix completion[J]. arXiv preprint arXiv:1706.02263, 2017.

[46] Liu Y, Xia L, Huang C. SelfGNN: Self-supervised graph neural networks for sequential recommendation[C]//Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2024: 1609-1618.

[47] Yang Y, Huang C, Xia L, et al. Debiased contrastive learning for sequential recommendation[C]//Proceedings of the ACM web conference 2023. 2023: 1063-1073.