



EMPNet: An extract-map-predict neural network architecture for cross-domain recommendation

Jinpeng Chen^{1,2,3,4} · Fan Zhang^{1,3} · Huan Li⁵ · Hua Lu⁶ · Xiongnan Jin⁷ ·
Kuien Liu^{8,9} · Hongjun Li² · Yongheng Wang⁷

Received: 17 November 2023 / Revised: 12 December 2023 / Accepted: 28 December 2023 /
Published online: 3 February 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Cross-domain recommendation leverages a user's historical interactions in the auxiliary domain to suggest items within the target domain, particularly for cold-start users with no prior activity in the target domain. Existing cross-domain recommendation models often overlook key aspects such as the complexities of transferring user interests between domains and the biases inherent in user behavior patterns. In contrast, our Extract-Map-Predict Neural Network Architecture (EMPNet) employs a disentanglement approach to map fine-grained user interests and utilize the biases inherent in the cross-domain recommendation. In feature extraction, we use the Bidirectional Encoder Representations from Transformers (BERT) and Identity-Enhanced Multi-Head Attention Mechanism to obtain the user and item feature vectors. In cross-domain user mapping, we disentangle the user feature vector into domain-shared and domain-specific interests for fine-grained cross-domain mapping to obtain the feature vector of cold-start users in the target domain. In rating prediction, we design a

This article belongs to the Topical Collection: *Special Issue on Advancing Recommendation Systems with Foundation Models*

Guest Editors: Kai Zheng, Renhe Jiang, and Ryosuke Shibasaki.

✉ Jinpeng Chen
jpchen@bupt.edu.cn

- ¹ School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing 100876, China
- ² Key Laboratory of Deep-time Geography and Environment Reconstruction and Applications of Ministry of Natural Resources, Chengdu University of Technology, Chengdu 610059, China
- ³ Key Laboratory of Trustworthy Distributed Computing and Service (Beijing University of Posts and Telecommunications), Ministry of Education, Beijing 100876, China
- ⁴ Xiangjiang Laboratory, Changsha 410205, China
- ⁵ College of Computer Science and Technology, Zhejiang University, Hangzhou 310058, China
- ⁶ Department of People and Technology, Roskilde University, Roskilde 4000, Denmark
- ⁷ Zhejiang Lab, Hangzhou 311121, China
- ⁸ Academy of Cyber, Beijing 100846, China
- ⁹ Huanghe S & T University, Zhengzhou 450006, China

biased Attentional Factorization Machine (AFM) to utilize biases extracted from user and item features. We experimentally evaluate EMPNet on the Amazon dataset. The results show that it clearly outperforms the selected baselines.

Keywords Cross-domain recommendation · Multi-head attention mechanism · VAE · Disentangle · Rating prediction

1 Introduction

Recommender systems play a crucial role in modern online applications like music/video websites and e-commerce platforms, recommending potentially interesting items to users based on their historical behavior. Various recommendation models have been developed, including Collaborative Filtering (CF) [1–5], which relies on user-item ratings, and some models utilize textual reviews [6–9]. Recently, the attention mechanism has been included in neural networks [10–15] to optimize recommendation performance by discriminating the contributions of different data features.

However, these recommendation models all suffer from the data sparsity problem. Data sparsity refers to that users in general interact with a limited number of items only, which leads to relatively insufficient user information and reduces the accuracy of recommendations. Particularly for cold-start problems, where users lack prior interactions in a specific domain, recommendation models encounter increased difficulty in suggesting suitable items for these users.

To address these two problems, especially the cold-start problem, researchers have proposed Cross-Domain Recommendation [16–22]. Cross-domain recommendation recommends to a user the item in one domain (*target domain*) by learning and transferring the user's historical behavior from another domain (*auxiliary domain*). An example of music domain recommendation by referring to the movie domain is illustrated in Figure 1. Existing cross-domain recommendation approaches often consist of three steps.

First, get the feature vector representation of the users. Second, a mapping of overlapping users is established from the auxiliary domain (movie in Figure 1) to the target domain (music in the example). Third, based on the user mapping, the recommendation for a cold-start user in the target domain is achieved by transferring the user preferences from the auxiliary domain.

However, existing cross-domain recommendation approaches still face several challenges. The first is the incomplete utilization of review information. User reviews contain various information about users and items, and some models fail to make use of information-rich user reviews (e.g., [17]), or underutilize reviews (e.g., [18]). Different reviews are not independent but related to each other. For example, a user is particularly harsh and always makes habitual negative reviews. In the reviews for a certain item, it is important to distinguish and separately process the reviews from the particular user and those from other normal users. Another challenge is how to map the interest transformation between different domains. In cross-domain recommendation, the user's interest is bifurcated into two facets: domain-shared and domain-specific interests. While the former can offer benefits across different domains, the latter is relevant only within a single domain, and applying it directly to another domain may result in adverse 'negative transfer' effects. Unfortunately, common methodologies often neglect this aspect, resulting in a conflation of interests across domains, which can subsequently diminish the accuracy of recommendations. Although recent advancements (e.g., [23–27]) have introduced techniques to disentangle these interests, they primarily capitalize on domain-

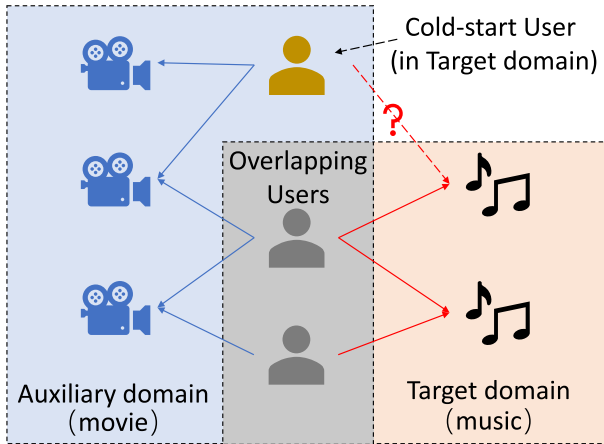


Figure 1 An example of cross-domain recommendation

shared interests to bolster overlapping users’ representations in both domains, often ignoring strategies for recommending to cold-start users.

Motivated as such, we propose an Extract-Map-Predict Neural Network Architecture (EMPNet) that exploits and differentiates user reviews for cross-domain recommendation.

EMPNet innovates and integrates multiple technologies to produce cross-domain recommendations. First of all, EMPNet considers the review information to increase data availability and diversity. When processing the review text information, EMPNet leverages Bidirectional Encoder Representations from Transformers (BERT) [28] and the Identity-Enhanced Multi-Head Attention Mechanism to improve the utilization of review information. Next, to improve the accuracy of the cross-domain user mapping, EMPNet employs a Domain Mapping Variational Autoencoder (DM-VAE). In particular, DM-VAE disentangles user feature vectors into domain-shared and domain-specific interests, transfers only the domain-specific interests across domains, and integrates transfer results with the domain-shared interests to get the feature vector of cold-start users in the target domain.

Finally, to improve the prediction accuracy of cross-domain recommendation, EMPNet improves an attentional factorization machine (AFM) by adding to it three biases that represent the inherent features of users, items, and domains. The user feature bias represents the user’s scoring habits, the item feature bias includes the intrinsic qualities of the item and the domain feature bias represents the overall rating of the domain.

We make the following major contributions:

- We propose EMPNet, an innovative model that discerns the significance of reviews and disentangles user interests to enhance cross-domain recommendation performance.
- We add the user ID to the review feature vector of multi-headed attention to distinguish users who write reviews of different quality and design a biased AFM for EMPNet by incorporating biases on user features, item features, and domain features to indicate their historical scoring preferences.
- We propose the DM-VAE method, which disentangles user feature vectors into domain-shared and domain-specific interests. It then transfers the domain-specific interests from the auxiliary domain to the target domain, where they are merged with the domain-shared interests to form the user feature vector in the target domain.

- We evaluate EMPNet on the Amazon dataset. The experimental results show that EMPNet clearly improves accuracy in the cross-domain recommendation.

The rest of the paper is organized as follows. Section 2 reviews related work. Section 3 elaborates on our EMPNet model. Section 4 reports on experiments. Section 5 concludes the paper and discusses future research.

2 Related work

Single-domain recommendation models roughly fall into two categories.

Review-based recommendation models utilize textual reviews on items to improve recommendations. ALFM (Aspect-aware Latent Factor Model) [6] models user preferences and item features on review texts, and integrates them with latent factors learned from the user-item rating matrix. MTPR (Multi-Task Pairwise Ranking) [8] combines collaborative embedding and content embedding to address the cold-start problem in the multimedia recommendation. For normal items, both kinds of embedding are used. For cold-start items, collaborative embedding is replaced by a zero vector. DeepCoNN (Deep Cooperative Neural Networks) [9] uses two parallel CNNs to generate representations of user behaviors and item properties over their related reviews, followed by a Factorization Machine (FM) to predict the user-item rating.

Attention-based recommendation models further discriminate the importance of different data items. AFM (Attentional FM) [13] weights all different feature interactions according to their contribution to the result. KGAT (Knowledge Graph Attention network) [14] embeds users, items, and item attributes from a knowledge graph. It uses an attention mechanism to compute the importance of graph neighbors. Note that none of the above-recommended models can resolve the cold-start problem.

Cross-domain recommendation models mitigate the cold-start problem by exploiting data from the auxiliary domain. π -Net (Parallel Information-sharing Network) [29] makes sequential recommendations simultaneously for two domains where the user behaviors are synchronously shared at each timestamp. Unlike π -Net, our EMPNet does not require shared accounts across domains. CPR (cross-domain paper recommendation) [30] is a cross-domain recommendation model for paper recommendation. CPR learns the probabilistic associations of paper content with the existing discipline classification. Then a user interest is represented as a probabilistic distribution over the target domain semantics. Finally, relevant papers are recommended to users according to user interest and paper content. EMCDR (Embedding and Mapping framework for Cross-Domain Recommendation) [17] captures non-linear mapping across different domains through an MLP embedding process. RC-DFM (Review and Content-based Deep Fusion Model) [18] uses additional stacked denoising autoencoder (aSDAE) [33] to fuse review texts and item contents with the rating matrix in both domains. CATN (Cross-domain recommendation framework via Aspect Transfer Network) [31] extracts multiple-aspect review documents as well as auxiliary reviews of users, and learns inter-aspect correlations across domains through an attention mechanism. PTUPCDR (Personalized Transfer of User Preferences for Cross-domain Recommendation) [32] learns a meta-network to generate personalized bridge functions to achieve personalized preference transfer for each user. In recent years articles applying disentangled representation learning in cross-domain recommendation have begun to emerge. SER [27] introduced a method that leverages user reviews for domain disentanglement, focusing on enhancing the performance of recommendation systems through text analysis and domain

Table 1 Feature comparison of recommendation models

Category	Model	Data Rating		Review		Algorithm	
		Rating	Disentangle	Review	Disentangle	Attention	Disentangle
Single-domain	ALFM [6]	✓		✓			
	MTPR [8]	✓		✓			
	DeepCoNN [9]	✓		✓			
Attention-based	AFM [13]	✓				✓	
	KGAT [14]	✓				✓	
	π -Net [29]	✓					
	CPR [30]			✓			
Cross-domain	EMCDR [17]	✓					
	RC-DFM [18]	✓		✓			
	CATN [31]	✓		✓			
	PTUPCDR [32]	✓				✓	
	SER [27]	✓		✓			✓
	DisenCDR [26]	✓					✓
	EMPNet(ours)	✓		✓		✓	✓

identification. However, this paper only focuses on cross-domain recommendations without overlapping users and does not utilize information from overlapping users. DisenCDR (Disentangled Representations for CDR) [26] learns disentangled user representations through mutual information regularizers to distinguish between domain sharing and specific information. This approach applies to the scenario of dual-domain boosting for shared users and does not apply to the cold-start problem.

Our EMPNet distinguishes itself from existing cross-domain recommendation approaches in two major aspects. First, we make full use of reviews and use an Identity-Enhanced Multi-Head Attention Mechanism to classify the importance of reviews. Second, most models directly map information from the auxiliary domain to the target domain without a finer delineation of representations across different domains. Some methods that employ disentangled representations have not fully harnessed the diverse information in cold-start scenarios, which involve data from both overlapping and cold-start users. We disentangle user representations and utilize overlapping users to learn cross-domain mappings for cold-start users.

Table 1 compares these Recommendation models.

3 EMPNet

Table 2 lists the symbols frequently used in this paper. We utilize overlapping users U_o in the auxiliary domain D^A and target domain D^T to make cross-domain recommendations for cold-start users U_c . We harness the textual content of reviews and associated review entities to construct feature vectors for both users and items. Specifically, an item's feature vector \mathbf{f}_i is generated from its aggregate historical reviews $C_i = \{c_{i1}, c_{i2}, \dots, c_{ik}\}$ and the corresponding users $U_i = \{u_{i1}, u_{i2}, \dots, u_{ik}\}$, while a user's feature vector \mathbf{f}_u is derived from their aggregate historical reviews C_u and the associated items I_u . The key of our approach is to effectively transfer the feature vectors of cold-start users from an auxiliary domain to the target domain, thereby improving accurate cross-domain recommendations.

Table 2 Notations

Symbol	Definition
$C_u (C_i)$	The historical reviews of user u (item i)
\mathbf{v}_{ik}	The review feature vector representation of the k -th review of the item i
\mathbf{o}_{ik}	The identity-enhanced review vector of the k -th review of the item i
$\mathbf{f}_u (\mathbf{f}_i)$	The feature vector of user u (item i)
U_i	A set of users who commented on item i
I_u	A set of items reviewed by user u
$b_u (b_i)$	Bias of user u (item i)
$b_{u_c} (b_{u_o})$	Bias of cold-start user u_c (overlapping user u_o)
e_u^A	domain-specific interests of user u in auxiliary domain
e_u^S	domain-shared interests of user u
\mathbf{R}^A	Rating matrix for users in auxiliary domain
r_{ui}	User u 's predicted rating on item i
$r_{u_c i}^T$	Cold-start user u_c 's predicted rating on item i

The architecture of EMPNet is shown in Figure 2, where symbols with the superscripts A and T represent data in auxiliary and target domains, respectively. The structure of EMPNet can be divided into three modules:

- The **feature extraction module** employs BERT and the Identity-Enhanced Multi-Head Attention Mechanism to extract user (resp. item) feature vectors \mathbf{f}_u (resp. \mathbf{f}_i) from review texts.
- The **cross-domain user mapping module** employs feature vectors from overlapping users \mathbf{f}_{u_o} to train the DM-VAE, learning the cross-domain mappings from the auxiliary to the target domain, which is subsequently applied to cold-start users to derive their feature vectors in the target domain $\mathbf{f}_{u_c}^T$. Additionally, this module utilizes an MLP to determine the bias $b_{u_c}^T$ for each of these cold-start users.
- The **prediction module** calculates the user u 's rating of each item by combining the user feature vector \mathbf{f}_u and the item feature vector \mathbf{f}_i . The item with the highest rating is recommended to the user u .

It is noteworthy that EMPNet supports both single-domain recommendation and cross-domain recommendation, and the former lies as the foundation for the latter. At first, EMPNet combines **feature extraction module** and **prediction module** to enable single-domain recommendation for each domain. For cross-domain recommendation, EMPNet takes the intermediate results of single-domain recommendations from the target and auxiliary domains, i.e., the feature vector of user \mathbf{f}_u and the feature vector of item \mathbf{f}_i , feeds it into **cross-domain user mapping module**, and finally runs another **prediction module** to realize cross-domain recommendation.

Next, we describe these three modules in detail.

3.1 Feature extraction module

Users typically produce reviews when purchasing items. Such reviews encompass both user and item feature data. To this end, this module extracts the corresponding feature vectors from these reviews. This module is common to both auxiliary and target domains. Also, the process of this module is the same for users as for items. For simplicity, Figure 2 only illustrates the process for items. The process for obtaining item feature vectors is as follows.

To extract features from reviews, this module uses BERT, a foundational model optimized for tasks in the realm of natural language processing. It can convert a sequence of text into a

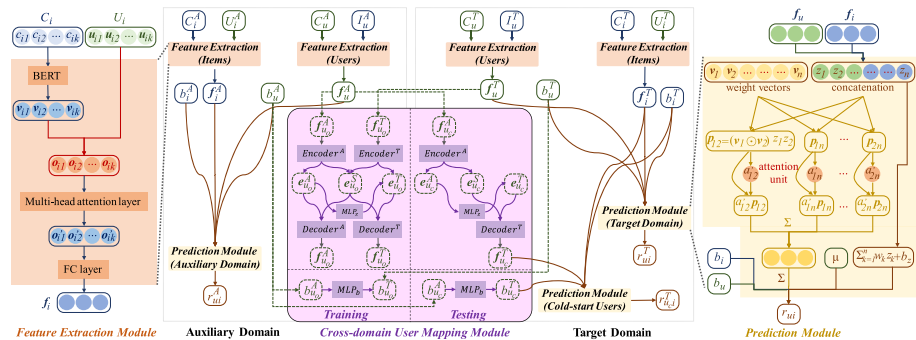


Figure 2 Overall architecture of EMPNet

fixed-size vector, capturing the contextual relationships between words. Given the review set C_i for an item i , this module converts each review c_{ik} into a review feature vector $\mathbf{v}_{ik} \in \mathbb{R}^{k_1}$ using BERT. Here, k_1 denotes the dimension of word vectors is 768 and the number of reviews is k_0 . The vector \mathbf{v}_{ik} encapsulates the sentiment and content of the review, making it a valuable feature for our recommendation system.

Every review corresponds to a user. Some users may tend to be more critical in their ratings, resulting in lower ratings, while others may have a stronger preference for certain items. To account for this, we introduce a user encoding $\mathbf{u}_{ik} \in \mathbb{R}^{k_1}$ for each review. This encoding is a complex vector representing the user’s profile.

As different reviews contribute differently to the overall evaluation of an item, this module assigns weights to reviews through a multi-head attention mechanism. The multi-head attention mechanism employed in this study is termed the Identity-Enhanced Multi-Head Attention Mechanism. According to the method of adding positional encoding to the input embeddings in a previous work [34], we add the user encoding \mathbf{u}_{ik} about who writes the review to the review feature vector \mathbf{v}_{ik} as $\mathbf{o}_{ik} \in \mathbb{R}^{k_1}$. The Identity-Enhanced Multi-Head Attention Mechanism is defined as:

$$\begin{aligned} \mathbf{o}_{ik} &= \mathbf{v}_{ik} + \mathbf{u}_{ik} & (1) \\ \mathbf{Q} = \mathbf{K} = \mathbf{V} = \mathbf{O}_i &= \{\mathbf{o}_{i1}, \mathbf{o}_{i2}, \dots, \mathbf{o}_{ik_0}\} \\ \text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_v}}\right)\mathbf{V} \\ \text{head}_i &= \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \\ \text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \text{Concat}(\text{head}_1, \dots, \text{head}_h)\mathbf{W}^O \\ \mathbf{O}'_i &= \{\mathbf{o}'_{i1}, \mathbf{o}'_{i2}, \dots, \mathbf{o}'_{ik_0}\} = \text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) \end{aligned}$$

where $\mathbf{O}_{ik} \in \mathbb{R}^{k_0 \times k_1}$ represents the set of reviews for item i , $\mathbf{Q} \in \mathbb{R}^{k_0 \times k_1}$, $\mathbf{K} \in \mathbb{R}^{k_0 \times k_1}$, $\mathbf{V} \in \mathbb{R}^{k_0 \times k_1}$ represents the query, keys, and values respectively, and the value of all three of them is the input vector \mathbf{O}_{ik} ; $\mathbf{o}'_{ik} \in \mathbb{R}^{k_1}$ represents the weighted \mathbf{o}_{ik} ; $\mathbf{W}_i^Q \in \mathbb{R}^{k_1 \times d_v}$, $\mathbf{W}_i^K \in \mathbb{R}^{k_1 \times d_v}$, $\mathbf{W}_i^V \in \mathbb{R}^{k_1 \times d_v}$, $\mathbf{W}^O \in \mathbb{R}^{hd_v \times k_1}$; h denote the number of heads respectively, and $d_v = k_1/h$.

After obtaining the weighted feature vector \mathbf{o}'_{ik} by the Identity-Enhanced Multi-Head Attention Mechanism, this module feeds the weighted sum of the review features \mathbf{f}'_i to the MLP to obtain the item feature vector \mathbf{f}_i . The \mathbf{f}_i is calculated as follows:

$$\begin{aligned} \mathbf{f}'_i &= \sum_{k=1}^{k_0} \mathbf{o}'_{ik} & (2) \\ \mathbf{f}_i &= f_{mlp}(\mathbf{f}'_i) \end{aligned}$$

where $\mathbf{f}'_i \in \mathbb{R}^{k_1}$ and $\mathbf{f}_i \in \mathbb{R}^{k_f}$. Here, k_f denotes the dimensionality of feature vectors.

For a user u , the user feature vector $\mathbf{f}_u \in \mathbb{R}^{k_f}$ is obtained in the same way, except that the input includes an item set I_u and review set C_u . The set I_u contains the items for which user u has written a review.

3.2 Prediction module

Then we use **prediction module** in both domains to accomplish single-domain recommendations. This module uses a biased AFM to predict a user’s rating for an item using their

feature vectors. The biased AFM consists of five parts: the paired interaction part, the linear regression part, the average rating of a domain, the user feature bias, and the item feature bias.

The paired interaction part works as follows. A pair of user and item feature vectors \mathbf{f}_u and \mathbf{f}_i are concatenated to generate the rating feature vector $\mathbf{z}_{ui} \in \mathbb{R}^n$, where n is the sum of the dimensionalities of \mathbf{f}_u and \mathbf{f}_i . The interaction result \mathbf{p}_{kl} between each pair of components z_k and z_l in \mathbf{z}_{ui} is calculated as

$$\mathbf{p}_{kl} = (\mathbf{v}_k \odot \mathbf{v}_l)z_kz_l, \quad z_k, z_l \in \mathbf{z}_{ui} \tag{3}$$

where \odot denotes the element-wise product of two vectors, $\mathbf{p}_{kl} \in \mathbb{R}^{k_2}$, $\mathbf{v}_k (\mathbf{v}_l) \in \mathbb{R}^{k_2}$ is the weight vector of $z_k (z_l)$, and k_2 denotes the dimensionality of the weight vector.

As the interaction result \mathbf{p}_{kl} does not always contribute to the final result with the same significance, we use an attention mechanism to get the attention score for a \mathbf{p}_{kl} .

$$\mathbf{a}'_{kl} = \mathbf{h}_p^T ReLU(\mathbf{W}_p \mathbf{p}_{kl} + \mathbf{b}_1) + b_2 \tag{4}$$

$$\mathbf{a}_{kl} = \frac{\exp(\mathbf{a}'_{kl})}{\sum_{k \in n, l \in n} \exp(\mathbf{a}'_{kl})} \tag{5}$$

where $\mathbf{W}_p \in \mathbb{R}^{k_3 \times k_2}$ is the weight matrix of \mathbf{p}_{kl} . We have $\mathbf{b}_1 \in \mathbb{R}^{k_3}$, $b_2 \in \mathbb{R}^1$ and $\mathbf{h}_p \in \mathbb{R}^{k_3}$. We normalize \mathbf{a}'_{kl} to \mathbf{a}_{kl} . The final result of the paired interaction part is obtained as

$$\mathbf{h}_p^T \sum_{k=1}^n \sum_{l=k+1}^n \mathbf{a}_{kl} (\mathbf{v}_k \odot \mathbf{v}_l) z_k z_l \tag{6}$$

where $\mathbf{h}_p^T \in \mathbb{R}^{k_3}$ is the weight of the paired interactive part.

The resultant formula of the biased AFM is

$$r_{ui} = \mathbf{h}_p^T \sum_{k=1}^n \sum_{l=k+1}^n \mathbf{a}_{kl} (\mathbf{v}_k \odot \mathbf{v}_l) z_k z_l + \left(\sum_{k=1}^n w_k z_k + b_z \right) + \mu + b_u + b_i \tag{7}$$

where r_{ui} is user u 's predicted rating on item i , and $w_k \in \mathbb{R}^1$ (resp., $b_z \in \mathbb{R}^1$) represents the weight (resp., bias) of the linear regression part. μ is the average rating of a domain which serves as the feature of that domain and can be calculated directly. b_u is the user feature bias indicating a user's scoring habits and b_i is the item feature bias representing the overall scoring situation of an item. Both b_u and b_i are subject to random initialization and undergo subsequent training to achieve optimal performance.

3.3 Cross-domain user mapping module

After the single-domain recommendation, we use its intermediate result: user feature vector and user bias for the cross-domain recommendation. To tackle the cold-start problem, we propose the cross-domain user mapping module.

In the auxiliary domain, user feature vectors \mathbf{f}_u^A are divided into those of overlapping users $\mathbf{f}_{u_o}^A$ and cold-start users $\mathbf{f}_{u_c}^A$. Likewise, user bias b_u^A is split into the bias of the overlapping user $b_{u_o}^A$ and the cold-start user $b_{u_c}^A$. In the target domain, user feature vectors \mathbf{f}_u^T and biases b_u^T are the overlapping user's feature vector $\mathbf{f}_{u_o}^T$ and bias $b_{u_o}^T$. This module aims to learn the user feature vector $\mathbf{f}_{u_c}^T$ and bias $b_{u_c}^T$ of a cold-start user in the target domain.

Map user feature vectors This module employs the DM-VAE approach to map cold-start user feature vectors \mathbf{f}_{u_c} from the auxiliary domain to the target domain. In either domain,

DM-VAE independently trains a VAE structure, which comprises an encoder and a decoder, both of which are constructed using MLP. As an example of an overlapping user on the auxiliary domain in Figure 2, we use the encoder to disentangle user feature vectors $\mathbf{f}_{u_o}^A$ into two sub-vectors, representing domain-shared interests $\mathbf{e}_{u_o}^S$ and domain-specific interests $\mathbf{e}_{u_o}^A$. The formula for this part is as follows:

$$\begin{aligned} \mu_1, \sigma_1, \mu_2, \sigma_2 &= \text{encoder}^A(\mathbf{f}_u^A) \\ \mathbf{e}_{u_o}^S &= \text{reparam}(\mu_1, \log(\sigma_1^2)) \\ \mathbf{e}_{u_o}^A &= \text{reparam}(\mu_2, \log(\sigma_2^2)) \\ \hat{\mathbf{f}}_u^A &= \text{decoder}^A(\mathbf{e}_{u_o}^S, \mathbf{e}_{u_o}^A) \end{aligned} \tag{8}$$

where the encoder first samples the input vector, generating two sets of means and variances $\mu_1, \sigma_1, \mu_2, \sigma_2$. Then, using the reparameterization trick, we generate samples from these means and variances, where one sample represents domain-shared interests $\mathbf{e}_{u_o}^S$ and the other represents domain-specific interests $\mathbf{e}_{u_o}^A$. These two interest vectors are then concatenated and processed through the decoder to retrieve the original feature vector.

To guarantee the effective disentanglement of domain-shared interests $\mathbf{e}_{u_o}^S$ and domain-specific interests $\mathbf{e}_{u_o}^A$, we utilize a pair of Kullback-Leibler (KL) divergence losses $\text{KL}_{\text{shared}}^A$ and $\text{KL}_{\text{specific}}^A$. Subsequently, a reconstruction loss $\mathcal{L}_{\text{recon}}^A$ is applied to ensure that the output of the decoder is a close approximation of the original input. Each loss is defined as:

$$\begin{aligned} \text{KL}_{\text{shared}}^A &= \text{KL}(\mu_1, \sigma_1^2) \\ \text{KL}_{\text{specific}}^A &= \text{KL}(\mu_2, \sigma_2^2) \\ \mathcal{L}_{\text{recon}}^A &= \text{loss}_{\text{recon}}(\mathbf{f}_u^A, \hat{\mathbf{f}}_u^A) \end{aligned} \tag{9}$$

After training a VAE in each of the two domains, we obtained the domain-specific interests of the overlapping users $\mathbf{e}_{u_o}^A$ and $\mathbf{e}_{u_o}^T$ in both domains as well as their domain-shared interests $\mathbf{e}_{u_o}^S$. Subsequently, we trained an MLP to understand how the domain-specific interests of users transition across different domains:

$$\mathbf{e}_{u_o}^T = f_{mlp}(\mathbf{e}_{u_o}^A) \tag{10}$$

We use a loss function \mathcal{L}_{map} to ensure that domain-specific interests are mapped from the auxiliary domain to the target domain. For the domain-shared interests, we employed a loss function \mathcal{L}_{com} to ensure that the domain-shared interests learned by the two VAEs are consistent.

Once trained with data from overlapping users, the DM-VAE can be applied to the feature vectors of cold-start users \mathbf{f}_{u_c} as shown in Figure 2. Since cold-start users only have interactions in the auxiliary domain, we first decompose them using the encoder of the auxiliary domain, obtaining their domain-specific interests $\mathbf{e}_{u_c}^A$ and domain-shared interests $\mathbf{e}_{u_c}^S$. Then, using the trained MLP, we map the domain-specific interests from the auxiliary domain $\mathbf{e}_{u_c}^A$ to the target domain $\mathbf{e}_{u_c}^T$. Finally, we obtain the domain-specific interests $\mathbf{e}_{u_c}^A$ and domain-shared interests $\mathbf{e}_{u_c}^S$ of the cold-start users in the target domain and use the decoder of the target domain to derive the feature vector of the cold-start users in the target domain $\mathbf{f}_{u_c}^T$.

Mapping user bias The user bias vector encapsulates the user’s intrinsic attributes. Given its relatively simple structure, we employ an MLP to learn the mapping of user bias from the auxiliary domain b_u^A to the target domain b_u^T , the loss is $\mathcal{L}_{\text{bias}}$. This is informed by the overlapping user bias b_{u_o} across the two domains. By applying this mapping approach to the

cold-start user's bias in the auxiliary domain $b_{u_c}^A$, we can deduce the cold-start user's bias in the target domain $b_{u_c}^T$.

Cross-domain recommendations Upon acquiring the feature vectors and biases for cold-start users in the target domain, we can deploy the prediction module to generate cross-domain recommendations. The distinctive aspect of applying the prediction module in the cross-domain scenario, as opposed to the single-domain scenario, lies in the utilization of data: user feature vector and bias use only the information of the cold-start user in the target domain, and the item feature vector and bias use the information of the items in the target domain that have already been trained.

After all item ratings are computed for cold-start users, the items with the highest ratings will be recommended to them.

3.4 Model training

For single-domain recommendations, the feature extraction module and prediction module are trained end-to-end. The loss function is the loss of predicted ratings.

For cross-domain recommendations, We use the intermediate results of single-domain recommendations as input and train the cross-domain user mapping module with feature vectors of overlapping users with bias. The loss function for cross-domain recommendation $\mathcal{L}_{\text{cross}}$ is:

$$\begin{aligned}\mathcal{L}_{\text{recon}} &= \mathcal{L}_{\text{recon}}^A + \mathcal{L}_{\text{recon}}^T & (11) \\ \mathcal{L}_{\text{KL}} &= \text{KL}_{\text{shared}}^A + \text{KL}_{\text{shared}}^T + \text{KL}_{\text{specific}}^A + \text{KL}_{\text{specific}}^T \\ \mathcal{L}_{\text{cross}} &= \alpha \mathcal{L}_{\text{recon}} + \beta \mathcal{L}_{\text{KL}} + \gamma (\mathcal{L}_{\text{map}} + \mathcal{L}_{\text{com}}) + \delta \mathcal{L}_{\text{bias}} + \epsilon \mathcal{L}_{\text{score}}\end{aligned}$$

where $\mathcal{L}_{\text{recon}}$ represents the reconstruction loss of VAE, \mathcal{L}_{KL} represents the KL scatter loss, $\mathcal{L}_{\text{score}}$ represents the loss of predicted ratings, and α , β , γ , δ and ϵ represent the weights of the components.

In our experiments, we use Adam [35] as the optimizer for training. It minimizes the error between the predicted rating and the real rating. We apply dropout to the review feature vector in the feature extraction module and the paired interactions part in the prediction module. We also apply L2 regularization to the weight matrices in the two attention mechanisms. These measures help to avoid overfitting.

4 Experiments

4.1 Dataset and evaluation metrics

The Amazon dataset¹ contains users, items, and ratings/reviews on items, where each rating is coupled with a review.

From the total 21 item categories, we select the three largest pairs of categories for experiments, namely movie-music, movie-book, and book-music. As some items and users receive only small numbers of reviews, we preprocess the data as follows. In particular, we select all items with more than 20 reviews in each domain, and then select the overlapping users with more than 10 reviews.

¹ <https://snap.stanford.edu/data/web-Amazon.html>

Table 3 Statistics of the datasets

	Movie	Music	Movie	Book	Book	Music
#(Users)	1,077		1,322		1,776	
#(Items)	1,568	1,002	2,108	1,050	859	1,199
#(Ratings)	54,581	31,798	80,460	35,508	19,960	28,696

Since excluding all other users affects the number of reviews on the selected items, we repeat the process. The statistics of the datasets are shown in Table 3.

Following [31, 32], we use Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) as the evaluation metrics of user-item rating prediction.

$$\text{RMSE} = \left(\frac{1}{H} \sum_{h=1}^H (r'_h - r_h)^2 \right)^{1/2} \quad (12)$$

$$\text{MAE} = \frac{1}{H} \sum_{h=1}^H |r'_h - r_h| \quad (13)$$

where H denotes the number of test ratings, with r'_h and r_h denoting the predicted and actual ratings for the h -th instance, respectively.

4.2 Experimental setup

We implement our framework using PyTorch and GPU. For each experiment, we randomly select half of the users and remove their information in the target domain, designating them as cold-start users. Initially, we conduct single-domain experiments in each domain, referred to as **EPNet** (Extract-Predict). For these single-domain experiments, we partition the data into training, validation, and test sets with a ratio of 8:1:1 and employ five-fold cross-validation to ensure the accuracy of the results. We utilize BERT for preprocessing review texts, with each input review limited to a maximum length of 512, resulting in an output review vector of 768 dimensions. Through grid search optimization, we set the learning rate, regularization parameter, number of multi-head attention heads, and dimensions of user/item feature vectors to 0.0005, 0.0001, 4, and 20, respectively. For the cross-domain recommendation parameters, we employ Bayesian optimization to determine the optimal parameters for each experimental group.

4.3 Cross-domain recommendation performance results

We compare EMPNet with the following four alternatives.

- EMCDR [17]: This model first applies matrix factorization to learn the latent factors, and then uses the MLP network to map the user latent factors from the auxiliary domain to the target domain.
- R-DFM [18]: It is a simple version of RC-DFM [18]². It merges ratings and reviews through the extended aSDAE to enhance the presentation of users and items. MLP is also adopted in cross-domain user mapping.

² We skip RC-DFM as it requires extra data input.

- CATN [31]: This model aims to extract multiple aspects from per-user and per-item review documents as well as auxiliary reviews of users with similar interests, and learn inter-aspect correlations across domains through an attention mechanism.
- PTUPCDR [32]: This model learns a meta-network fed by user feature embeddings to generate personalized bridge functions that achieve personalized preference transfer for each user.

For the evaluation metrics RMSE and MAE, the results are shown in Table 4. Clearly, EMPNet outperforms all baselines in most cross-domain recommendations, demonstrating the superiority of our proposed model. The result of EMCDDR is the worst. The main reason is that this model does not use reviews, and the use of ratings is relatively simple compared to other baselines. R-DFM incorporates review information as incidental content into the rating mechanism. This results in a low utilization rate of reviews, thus making the matrix factorization method adopted in predicting ratings ineffective. The CATN model extracts multiple aspects of users and items from the review documents for cross-domain transfer, making full use of the review data, so the performance is better than the R-DFM. On a majority of datasets, PTUPCDR demonstrates performance surpassed only by EMPNet, a distinction attributable to its innovative application of personalized preference transfer.

It is worth noting that on both the “Book to Music” and “Music to Book” experiments, the results of CATN are better than those of PTUPCDR, which may be attributed to the amount of data in the experiments. It can be seen from Table 3 that experiments with “Book to Music” and “Music to Book” have the least amount of data. The main reason is that the use of additional review data improves the results more significantly when the amount of data is small.

In the “Book to Movie” experiment, the performance improvement of EMPNet is relatively small, which is related to the particularity of the data. It can be seen from Table 3 that the “Book to Movie” experiment has the largest gap between the number of items in the auxiliary domain and the target domain. EMPNet performs the same operations on users and items in multiple steps, whereas CATN and PTUPCDR have a predilection for user-centric information extraction. Consequently, in the “Book to Movie” experiment, the superiority of

Table 4 Performance of cross-domain recommendation on RMSE and MAE

		EMCDDR	R-DFM	CATN	PTU-PCDR	EMPNet	Improve
Movie	RMSE	1.4942	1.3216	1.0591	<u>1.0087</u>	0.8593	+14.81%
to Music	MAE	1.1891	0.9628	0.7852	<u>0.7505</u>	0.7426	+1.05%
Music	RMSE	1.5411	1.3621	1.0792	<u>1.0558</u>	0.9036	+14.42%
to Movie	MAE	1.233	1.0301	0.8456	<u>0.8052</u>	0.8004	+0.60%
Movie	RMSE	1.5199	1.2265	0.9870	<u>0.9810</u>	0.8145	+16.97%
to Book	MAE	1.2223	0.9176	0.7645	<u>0.7465</u>	0.7293	+2.30%
Book	RMSE	1.3659	1.2693	0.9877	<u>0.9827</u>	0.8629	+12.19%
to Movie	MAE	1.0981	0.9539	0.7611	0.7493	<u>0.7554</u>	-0.81%
Book	RMSE	1.6082	1.1357	<u>1.0035</u>	1.0664	0.7799	+22.27%
to Music	MAE	1.2666	0.8451	<u>0.7218</u>	0.7961	0.6947	+3.75%
Music	RMSE	1.6756	1.2180	<u>1.0470</u>	1.0764	0.8600	+17.86%
to Book	MAE	1.3499	0.9288	<u>0.8044</u>	0.8111	0.7662	+4.75%

The best and second-best results are highlighted in boldface and underlined respectively

our EMPNet is somewhat subdued. Conversely, in the “Book to Music” experiment, the gap in the number of items is the smallest, which directly reflects EMPNet’s most pronounced performance improvement in this experiment.

While EMPNet demonstrates significant improvements in RMSE, the enhancements in MAE are not so pronounced. This discrepancy may stem from RMSE’s strengthened sensitivity to larger prediction errors, which our model’s optimization strategy may be able to mitigate more effectively. Given that our optimization function is tailored for RMSE, this could also account for the less noticeable performance gains in reducing average errors compared to squared errors. We leave it for future work to explore the adoption of alternative loss functions to achieve a more balanced enhancement across both metrics.

4.4 Ablation study

As mentioned in Section 4.3, our proposed model outperforms the baselines. These improvements come from three innovations of our model: Identity-Enhanced Multi-Head Attention Mechanism in the feature extraction module, DM-VAE in the cross-domain user mapping module, and biased AFM in the prediction module.

In this section, we conduct an ablation study to demonstrate the importance of each of the three innovations. Given that the feature extraction module and the prediction module can constitute a single-domain recommendation, we directly test the efficacy of the Identity-Enhanced Multi-Head Attention Mechanism and the biased AFM in the single-domain recommendation. The effectiveness of DM-VAE is evaluated within the cross-domain recommendation. Specifically, we compare the proposed model with the following variants:

- EPNet-ATN: It does not utilize the identity information of the reviews. Instead, it directly feeds the review vectors output by BERT into the multi-head attention mechanism in the feature extraction module.
- EPNet-AFM: It replaces the biased AFM with the unbiased AFM in the prediction module.
- EMPNet-MLP: It uses the ordinary MLP without DM-VAE in the cross-domain user mapping module of EMPNet.

The results of the ablation experiments are shown in Tables 5 and 6, and our proposed designs are effective in all experiments. The results show that using the Identity-Enhanced Multi-Head Attention Mechanism improves the results, and adding the user or item information corresponding to the review is beneficial to identifying valuable reviews. The performance results also verify the effectiveness of the biased AFM. Without the biases representing inherent features of users, items, and domains, EPNet-AFM makes less relevant recommendations. Also, EMPNet outperforms EMPNet-MLP. This can be attributed to the

Table 5 Performance of the Ablation Study on EPNet-MLP and EPNet-AFM (best result in bold)

		EPNet-ATN	EPNet-AFM	EPNet
Music domain	RMSE	0.8582	0.8670	0.8482
	MAE	0.7789	0.7871	0.7676
Movie domain	RMSE	0.9623	0.9780	0.8856
	MAE	0.8675	0.8886	0.7957
Book domain	RMSE	0.8903	0.9499	0.8835
	MAE	0.8123	0.8638	0.7975

Table 6 Performance of the ablation study on EMPNet-MLP (best result in bold)

		EMPNet-MLP	EMPNet
Movie to Music	RMSE	0.8766	0.8593
	MAE	0.7848	0.7426
Music to Movie	RMSE	0.9039	0.9036
	MAE	0.8113	0.8004
Movie to Book	RMSE	0.8146	0.8145
	MAE	0.7298	0.7293
Book to Movie	RMSE	0.9252	0.8629
	MAE	0.8371	0.7554
Book to Music	RMSE	0.7800	0.7799
	MAE	0.6960	0.6947
Music to Book	RMSE	0.8607	0.8600
	MAE	0.7724	0.7662

DM-VAE disentangling the user feature vector into domain-shared interests and domain-specific interests. By the strategy of only mapping domain-specific interests while keeping domain-shared interests unchanged, it can realize more accurate cross-domain interest transfer compared to directly mapping the entire user feature vector.

5 Conclusion and future work

In this paper, we propose a cross-domain recommendation model EMPNet. For feature extraction, EMPNet uses the BERT and Identity-Enhanced Multi-Head Attention Mechanism to distinguish the impact of different quality user and item reviews on the ratings. For cross-domain user mapping, EMPNet employs DM-VAE to disentangle the user feature vector into domain-shared and domain-specific interests, facilitating the cross-domain transfer to derive the cold-start user's feature vector in the target domain. For rating prediction, EMPNet considers and differentiates multiple kinds of biases that represent the inherent features of users, items, and domains. Experiments on real data verify the effectiveness of these designs and the performance superiority of EMPNet.

Several directions exist for future work. First, input data from multiple, diverse auxiliary domains may further improve cross-domain recommendation. Second, combining conventional recommendation models with foundation models may help cross-domain recommendation. Third, using multi-modal reviews, such as image and video may also improve cross-domain recommendation.

Author Contributions Jinpeng Chen wrote the main manuscript text and provided the methodology and funding support. Fan Zhang carried out the experiments. Huan Li and Hua Lu conceived the study and participated in methodology design and coordination. Xiongnan Jin and Kuien Liu helped draft the manuscript. Hongjun Li and Yongheng Wang provided writing review and editing and prepared all figures and tables.

Funding This work was supported in part by Beijing Natural Science Foundation(Grant No.L233034), in part by Zhejiang Lab Open Research Project (Grant No.K2022KG0AB03), in part by the National Natural Science Foundation of China (Grant No.62306287, No.62002027, No.62006023), in part by the National Key R&D Program of China (Grant No.2022YFE0137800), in part by Open Fund (DGERA 20231101) of Key Laboratory of Deep-time Geography and Environment Reconstruction and Applications of Ministry of Natural Resources, Chengdu University of Technology, in part by Zhejiang Provincial Natural Science Foundation of China (Grant

No.LY23F020012), in part by CCF-Zhipu AI Large Model Fund (Grant No. CCF-Zhipu202317), in part by SMP-IDATA Open Youth Fund (No.SMP2023-iData-005), in part by the Fundamental Research Funds for the Central Universities (Grant No.2023RC08, No.21623402), in part by the Open Project of Xiangjiang Laboratory (No.23XJ03006) and in part by the Open Projects of the Technology Innovation Center of Cultural Tourism Big Data of Hebei Province (Grant No.SG2019036-zd202205).

Availability of data and materials The data and materials are available from the corresponding author upon reasonable request.

Declarations

Ethics approval Not applicable

Competing interests The authors have no conflict of interest.

References

1. Yu, R., Ye, D., Wang, Z., Zhang, B., Move, O.A., Li, J., Jin, B., Kurdahi, F.J.: CFFNN: cross feature fusion neural network for collaborative filtering. *IEEE Trans. Knowl. Data Eng.* **34**(10), 4650–4662 (2022)
2. Wang, W., Tang, T., Xia, F., Gong, Z., Chen, Z., Liu, H.: Collaborative filtering with network representation learning for citation recommendation. *IEEE Trans. Big Data.* **8**(5), 1233–1246 (2022)
3. Yang, M., Li, Z., Zhou, M., Liu, J., King, I.: HICF: hyperbolic informative collaborative filtering. In: *KDD '22: The 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Washington, DC, USA, August 14 - 18, 2022, pp. 2212–2221 (2022)
4. Lian, D., Chen, J., Zheng, K., Chen, E., Zhou, X.: Ranking-based implicit regularization for one-class collaborative filtering. *IEEE Trans. Knowl. Data Eng.* **34**(12), 5951–5963 (2022)
5. Long, J., Chen, T., Nguyen, Q.V.H., Xu, G., Zheng, K., Yin, H.: Model-agnostic decentralized collaborative learning for on-device POI recommendation. In: *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023*, Taipei, Taiwan, July 23–27, 2023, pp. 423–432 (2023)
6. Cheng, Z., Ding, Y., Zhu, L., Kankanhalli, M.: Aspect-aware latent factor model: rating prediction with ratings and reviews. In: *WWW*, pp. 639–648 (2018)
7. Sun, R., Cao, X., Zhao, Y., Wan, J., Zhou, K., Zhang, F., Wang, Z., Zheng, K.: Multi-modal knowledge graphs for recommender systems. In: *CIKM'20: The 29th ACM International Conference on Information and Knowledge Management, Virtual Event, Ireland, October 19–23, 2020*, pp. 1405–1414 (2020)
8. Du, X., Wang, X., He, X., Li, Z., Tang, J., Chua, T.: How to learn item representation for cold-start multimedia recommendation? In: *MM*, pp. 3469–3477 (2020)
9. Zheng, L., Noroozi, V., Yu, P.S.: Joint deep modeling of users and items using reviews for recommendation. In: *WSDM*, pp. 425–434 (2017)
10. Ni, J., Huang, Z., Yu, C., Lv, D., Wang, C.: Comparative convolutional dynamic multi-attention recommendation model. *IEEE Trans. Neural Networks Learn. Syst.* **33**(8), 3510–3521 (2022)
11. Wang, H., Liu, G., Liu, A., Li, Z., Zheng, K.: DMRAN: A hierarchical fine-grained attention-based network for recommendation. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019*, Macao, China, August 10–16, 2019, pp. 3698–3704 (2019)
12. Xie, J., Cui, Y., Huang, F., Liu, C., Zheng, K.: MARINA: an mlp-attention model for multivariate time-series analysis. In: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17–21, 2022*, pp. 2230–2239 (2022)
13. Xiao, J., Ye, H., He, X., Zhang, H., Wu, F., Chua, T.: Attentional factorization machines: learning the weight of feature interactions via attention networks. In: *IJCAI*, pp. 3119–3125 (2017)
14. Wang, X., He, X., Cao, Y., Liu, M., Chua, T.-S.: Kgat: knowledge graph attention network for recommendation. In: *KDD*, pp. 950–958 (2019)
15. Du, X., Yuan, H., Zhao, P., Qu, J., Zhuang, F., Liu, G., Liu, Y., Sheng, V.S.: Frequency enhanced hybrid attention network for sequential recommendation. In: *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2023*, Taipei, Taiwan, July 23–27, 2023, pp. 78–88 (2023)
16. Wang, C., Niepert, M., Li, H.: Recsys-dan: discriminative adversarial networks for cross-domain recommender systems. *IEEE Trans. Neural Networks Learn. Syst.* **31**(8), 2731–2740 (2020)

17. Man, T., Shen, H., Jin, X., Cheng, X.: Cross-domain recommendation: an embedding and mapping approach. In: IJCAI, pp. 2464–2470 (2017)
18. Fu, W., Peng, Z., Wang, S., Xu, Y., Li, J.: Deeply fusing reviews and contents for cold start users in cross-domain recommendation systems. In: AAAI, vol. 33, pp. 94–101 (2019)
19. Xu, J., Song, J., Sang, Y., Yin, L.: CDAML: a cluster-based domain adaptive meta-learning model for cross domain recommendation. *World Wide Web (WWW)*, **26**(3), 989–1003 (2023)
20. Li, P., Tuzhilin, A.: Dual metric learning for effective and efficient cross-domain recommendations. *IEEE Trans. Knowl. Data Eng.* **35**(1), 321–334 (2023)
21. Liu, J., Huang, W., Li, T., Ji, S., Zhang, J.: Cross-domain knowledge graph chiasmal embedding for multi-domain item-item recommendation. *IEEE Trans. Knowl. Data Eng.* **35**(5), 4621–4633 (2023)
22. Zhang, T., Chen, C., Wang, D., Guo, J., Song, B.: A vae-based user preference learning and transfer framework for cross-domain recommendation. *IEEE Trans. Knowl. Data Eng.* **35**(10), 10383–10396 (2023)
23. Zhu, J., Wang, Y., Zhu, F., Sun, Z.: Domain disentanglement with interpolative data augmentation for dual-target cross-domain recommendation. In: Proceedings of the 17th ACM Conference on Recommender Systems, RecSys 2023, Singapore, September 18–22, 2023, pp. 515–527 (2023)
24. Guo, X., Li, S., Guo, N., Cao, J., Liu, X., Ma, Q., Gan, R., Zhao, Y.: Disentangled representations learning for multi-target cross-domain recommendation. *ACM Trans. Inf. Syst.* **41**(4), 85–18527 (2023)
25. Zhang, R., Zang, T., Zhu, Y., Wang, C., Wang, K., Yu, J.: Disentangled contrastive learning for cross-domain recommendation. In: Database Systems for Advanced Applications - 28th International Conference, DASFAA 2023, Tianjin, China, April 17–20, 2023, Proceedings, Part II. Lecture Notes in Computer Science, vol. 13944, pp. 163–178 (2023)
26. Cao, J., Lin, X., Cong, X., Ya, J., Liu, T., Wang, B.: Disencdr: learning disentangled representations for cross-domain recommendation. In: SIGIR '22: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, July 11 - 15, 2022, pp. 267–277 (2022)
27. Choi, Y., Choi, J., Ko, T., Byun, H., Kim, C.: Review-based domain disentanglement without duplicate users or contexts for cross-domain recommendation. In: Proceedings of the 31st ACM International Conference on Information & Knowledge Management, Atlanta, GA, USA, October 17–21, 2022, pp. 293–303 (2022)
28. Devlin, J., Chang, M., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*. (2018). [arXiv:1810.04805](https://arxiv.org/abs/1810.04805)
29. Ma, M., Ren, P., Lin, Y., Chen, Z., Ma, J., Rijke, M.d.: π -net: a parallel information-sharing network for shared-account cross-domain sequential recommendations. In: SIGIR, pp. 685–694 (2019)
30. Xie, Y., Sun, Y., Bertino, E.: Learning domain semantics and cross-domain correlations for paper recommendation. In: SIGIR, pp. 706–715 (2021)
31. Zhao, C., Li, C., Xiao, R., Deng, H., Sun, A.: CATN: cross-domain recommendation for cold-start users via aspect transfer network. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25–30, 2020, pp. 229–238 (2020)
32. Zhu, Y., Tang, Z., Liu, Y., Zhuang, F., Xie, R., Zhang, X., Lin, L., He, Q.: Personalized transfer of user preferences for cross-domain recommendation. In: WSDM '22: The Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event / Tempe, AZ, USA, February 21 - 25, 2022, pp. 1507–1515 (2022)
33. Dong, X., Yu, L., Wu, Z., Sun, Y., Yuan, L., Zhang, F.: A hybrid collaborative filtering model with deep structure for recommender systems. In: AAAI, pp. 1309–1315 (2017)
34. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4–9, 2017, Long Beach, CA, USA, pp. 5998–6008 (2017)
35. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. In: ICLR (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.