



Heterogeneous graph convolutional network pre-training as side information for improving recommendation

Phuc Do¹ · Phu Pham²

Received: 11 September 2021 / Accepted: 29 March 2022 / Published online: 2 May 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract

For the recommendation domain, most of the existing integrated graph neural network (GNN)-based architectures have still much focused on encoding the associated extra side information in forms of heterogeneous information network (HIN). Then, it is simultaneously utilized in multiple fine-tuning processes to effectively learn preferences from user–item interaction data which might require tremendous computational efforts. In addition, these approaches have also failed to incorporate with previous learnt and transferable knowledge from pre-trained models to better fine-tune for recommendation task. To meet these challenges, in this paper we propose a novel heterogeneous graph neural architecture, named: PreHIN4Rec. The proposed PreHIN4Rec is considered as the graph pre-training approach for leveraging the performance of recommendation task in both accuracy and scalability aspects. In general, our proposed PreHIN4Rec is designed to efficiently preserve both heterogeneous schematic and local structural latent features of user–item interactions in forms of HINs. It can effectively support to better fine-tune for achieving remarkable improvements in recommendation tasks through integrating with existing recommendation frameworks. The extensive experiments in benchmark datasets demonstrate the effectiveness of our proposed PreHIN4Rec model in comparing with recent state-of-the-art GNN-based recommendation baselines.

Keywords Recommendation · Pre-train · GNN · Heterogeneous network

1 Introduction

In recent years, with the tremendous growths of the Internet there are emerges of massive social media platforms such as e-commerce and social network. These massive social networks [1, 2] have led to the increases in the importance of personalized recommendation systems (a.k.a. recommender system) [3–7]. In general, recommendation system can efficiently support users to personally reach out their actual interesting items (e.g. products, online news, services, etc.) as well as eliminate non-relevant items/information to maximize users' contentment. These recommendation systems have supported to significantly

benefit both customer's experiences and businesses. Within the explosive raises of information available on online social media platforms, users are frequently greeted with massive and countless products in e-commerce platforms (e.g. Amazon, Alibaba, etc.), movies, music, etc., in social media/entertainment platforms (Netflix, YouTube, etc.). Thus, the recommendation systems are necessary to eliminate problems related to information overload or customer over-choice. Traditionally, most of the recommendation techniques are designed to learn the preferences of users and items from their historical interactions in order to facilitate the user's utility prediction process on his/her unseen items. In general, most of the traditional recommendation systems can be categorized into two trends, which are the content-based and collaborative filtering-based approaches. These both recommendation approaches are designed to focus on different aspects of user–item interactions to assist the user and item characterization process and capture the personalized users' interests of users and items' attractiveness. However, these classical

✉ Phuc Do
phucdo@uit.edu.vn; pta.phu@hutech.edu.vn

¹ University of Information Technology (UIT), Viet Nam National University, Ho Chi Minh City, Vietnam

² Faculty of Information Technology, HUTECH University, Ho Chi Minh City, Vietnam

recommendation techniques in both approaches still suffered major challenges [4, 8] regarding with the simplicity, noise/sparsity in dataset as well as the cold-start problem.

In recent years, thanks to the rapid developments of computer's architecture as well as deep learning approach in multiple disciplines, the modern recommendation systems have been shifted to another level of better user–item interaction understating as well as representation learning. The recent advanced deep learning architecture has been emerged as the powerful and state-of-the-art representation learning solutions. Recent proposed deep learning-based models enable to effectively preserve the complex structural datasets and to be trained to achieve significant performances in multiple data-driven analysis and filtering tasks, such as recommendation. The learnt rich structural and contextual latent representations of users and items via deep learning-based techniques can be utilized to accurately predict users' utilities on non-interacted items. Thus, they can sufficiently alleviate challenges which are related to the data sparsity and cold-start problems. In general, the deep learning-based recommendation architectures are naturally considered as the collaborative user–item interaction representation learning approach. In this approach, the rich semantic embedding matrices of users and items are learnt and fine-tuned for handling the user's utility prediction process. However, these recent deep learning-based recommendation models still concentrated much on evaluating the homogeneous relationships user–item interaction. Moreover, they might also be unable to incorporate with extra auxiliary information associated with users and items. In recent times, the heterogeneous information network (HIN) analysis and mining [9–11] is considered as a potential direction for modelling complex structural recommendation datasets with multiple types of nodes and relationships. The principles of HIN have been widely studied in multiple primitive networked data analysis and mining tasks, such as similarity search [9, 11, 12], link prediction [13], community detection [14]. The heterogeneous side information resources are tightly associated with user and item nodes in forms of heterogeneous networks. Therefore, the adoption of HIN and GNN-based representation learning as side information in recommendation is widely attended and has become a main stream for most of recent studies in this area.

1.1 Recent achievements and challenges in recommendation

Information network embedding for recommendation with GNN For many years, with the dramatic progresses of deep learning, specifically with GNN-based architectures (e.g. graph convolutional network (GCN) [15], GraphSage [16],

graph attention network (GAT) [17]) and graph-based structural representation learning, there are multiple powerful techniques and recommendation-aware methods with GNN have been proposed. These advanced deep learning-based architectures have demonstrated significant improvements in multiple downstream recommendation tasks, such as the well-known NCF [6], NGCF [7] and LightGCN [18] which utilized the GCN-based architecture in order to leverage the multi-level ordered contextual representation learning between user and item relationships. These rich structural user and item representations are later used to enhance the recommendation performance by exploiting the cross-contextual interactive paths between user and item nodes. However, this network embedding-based approach still considered the user–item relationships as the homogeneous direct interactions. Thus, they totally failed to integrate with the extra associated information to achieve better recommendation outputs. In the context of real-world applications and actual recommendation datasets, the homogeneous user–item interaction data are highly sparse and might contain a lot of noises. Therefore, it is necessary to incorporate the current recommendation systems to extra rich side information (e.g. brands/categories/textual descriptions of products, ages/genders/occupations of users, etc.) in order to address these noise/sparsity related problems. Since, there are many studies [19–22] have presented advantages of modelling and analysing user–item interaction data as well as their associated information in forms of heterogeneous information networks. This integrated HIN-based approach enables to deeply understand as well as sufficiency learn the cross-contextual user and item insight knowledge to further improve the recommendation performance.

Heterogeneous network as side/exogenous information for recommendation task Mainly relying on the ideas of utilizing HIN-based data as side information resources, many researchers have attempted to efficiently integrate and model the user–item interactions as heterogeneous networks which enables to enrich the recommendation context. Recently, there are several notable studies have been proposed, such as: FMG [19], HERec [20], HetNERec [23], MAGNN [21] and MetaHIN [22]. These integrated HIN-based models have utilized the meta-path/meta-graph-based networked data representation learning techniques to effectively capture the rich heterogeneous relationships between user and item nodes. The learnt heterogeneous information sources are considered as side information to facilitate the recommendation process. By efficiently modelling and learning the heterogeneous rich contextual information between interactive semantic paths (in forms of meta-paths/meta-graphs [24]) between users and items, the integrated HIN and deep learning-based models have

shown remarkable improvements in recommendation task. Thus, the combination between heterogeneous network analysis and deep learning-based network embedding techniques is considered as a promising direction for achieving further unlimited enhancements in multiple tasks of recommendation domain.

Existing drawbacks of deep learning-based recommendation techniques However, most of recent deep learning-based recommendation systems over HIN (RSHIN) is considered as a high time-consuming approach. In fact, these models require tremendous computational efforts as well as times to simultaneously learn the complex structure of a given HIN. Then, these rich heterogeneous structural representations of users and items are later fine-tuned for multiple tasks to achieve the reasonable trade-off solutions for both HIN-based representation learning and recommendation tasks. In addition, the contemporary proposed GNN-based and RSHIN-based models still be unable to produce stable recommendation outputs in different datasets with different initializations of model's parameters. Moreover, they also lacked the capability of incorporating with previous pre-trained graph-based structural representations of users and items. Within the transfer learning context, these representation learning resources are normally considered as the transferable knowledge resources [25, 26]. Thus, it is vital to find a better solution for both model's parameter initialization and GNN-based pre-train model integration problems in the recommendation area. Recently, there is a notable work of Meng et al. [25] in proposing an GNN-based pre-trained network embedding approach for improving the performance of recommendation problem, called as: GCN-P/COM-P [25]. However, they still concentrated much on exploiting the homogeneous user–item interaction relationships as well as didn't explicitly integrate the pre-training GNN-based model with the user and item associated auxiliary information in the form of a heterogeneous network.

1.2 Our motivations and contributions

Motivations from existing gaps in recommendation Motivated by recent achievements in the pre-trained representation learning paradigm [25] of computer vision/natural language processing and heterogeneous network embedding domains [27], in this paper, we proposed a novel heterogeneous GNN-based pre-training schema for recommendation, called as PreHIN4Rec. Our proposed pre-trained network embedding-based recommendation model is designed to deal with aforementioned challenges (as illustrated in Fig. 1). In general, our proposed PreHIN4Rec is defined as a general graph neural pre-training model for effectively recommendation-driven preserving the user and

item interaction relationships in form of network heterogeneity. Then, the achieved rich contextual user and item embedding matrices are utilized for the continuous fine-tuning process for dealing with user's utility prediction. To do this, we integrate our model with different recent well-known recommendation frameworks, such as: matrix factorization (MF) [5], neural collaborative filtering (NCF) [6], NGCF [7] and LightGCN [18].

A transferrable network representation learning approach for recommendation Specifically, a multi-layered GCN-based architecture is applied in our proposed graph-based pre-training model to efficiently explore the heterogeneous multi-typed nodes and relationships of a given graph-structured recommendation data in forms of multi-relational sub-graph representation learning approach which is majorly inspired in previous works [25]. In order words, the previous rich schematic heterogeneous network representation learning is pre-trained and later used to facilitate the recommendation process in the context of network user–item interaction evaluation. As targeting for handling recommendation task, we force the GCN-based pre-training model to focus on capturing semantic interactive paths between user and item node type. These user–item interactive semantic paths are represented as different meta-paths which convey various rich schematic meanings between user–item interaction relations. These multiple meta-path-based embeddings of user and item nodes are fused with multi-relational embeddings and then jointly optimized for the recommendation-driven learning process. By incorporating with the meta-path-based representation learning strategy, our heterogeneous GNN-based pre-training model can efficiently encode the rich heterogeneous schematic representations of users and items which are later applied to fine-tune for any general recommendation task as the pre-trained transferrable knowledge resource.

Our main contributions In general, our contributions in this paper with the corresponding proposed PreHIN4Rec model can be summarized as three-folds, which are:

- First of all, we formally introduce the utilization of heterogeneous GCN-based pre-training paradigm for preserving multi-level ordered and rich schematic user and item node representations. To do this, in our approach, we utilize the combination between multi-relational sub-graph [25] and meta-path-based embedding strategies [28, 29] within the context of HIN-based representation learning. These archived user and item representations are then fine-tuned as an end-to-end neural learning architecture to produce a completed GNN-based pre-trained model which are aimed to handle the general recommendation-driven purpose.

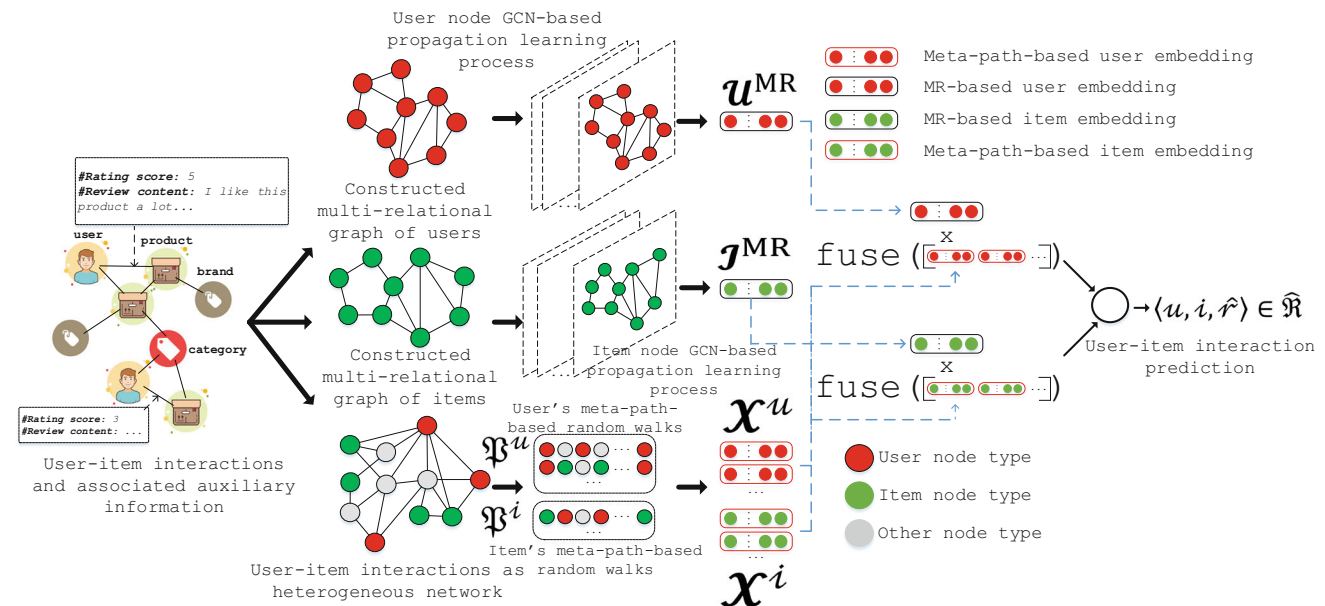


Fig. 1 The illustration of overall our proposed PreHIN4Rec pre-training schema for recommendation task

- Next, within the given pre-training GCN-based architecture, in order to efficiently transform the separated multi-relational and multiple meta-path-based user and item representations into the unified embedding space, we implement a neural fusion mechanism. This neural fusion mechanism supports to softly characterize and merge different user and item embeddings. They effectively support to fulfil both heterogeneous side information modelling and representation learning objectives. In the end, all the fusion mechanism's parameters are jointly optimized with other model's parameters during the learning process.
- Finally, the achieved heterogeneous GCN-based pre-trained model are deployed as the graph-structured transferable knowledge resources. Then, these pre-trained structural knowledge resources are fine-tuned with several existing notable recommendation frameworks, such as: MF [5], Metapath2Vec [29], NCF [6], NGCF [7], LightGCN [18] and SGL [30] to remarkably improve the performance of recommendation task. The extensive experiments in benchmark datasets, such as MovieLens, Foursquare and Amazon-Book for the top-n item recommendation task demonstrated the effectiveness of our proposed PreHIN4Rec model in comparing with recent state-of-the-art recommendation baselines. Moreover, we also conducted thorough ablation studies to show the usefulness of using pre-trained user and item embeddings for significantly improving recommendation outputs in both accuracy and stability performance aspects.

In overall, the left contents of our paper are organized into 4 sections. In the next section, we briefly present our recent works which are related to the deep learning-based recommendation domain as well as discuss about pros/cons of existing models. In the third section, we formally describe the methodology and implementation of our proposed PreHIN4Rec model. Next, we show extensive experiments as well as ablation studies on benchmark datasets to evaluate the performance of our proposed model in comparing with recent recommendation baselines. Finally, we conclude our achievements in this paper and provide some potential directions for future works in the fifth section.

2 Related works

In recent years, the tremendous development of deep learning-based approaches in multiple disciplines have led to significant progresses in multiple data-driven analysis and learning tasks, including recommendation [1, 2, 6, 7]. For a long run, the recommendation system has been considered as an indispensable application. It supports to alleviate the information overload/customer over-choice upon multiple large-scale online social (Facebook, Twitter, Baidu, etc.)/e-commercial (Amazon, Alibaba, etc.) platforms. Back to few decades, multiple classical recommendation approaches have been proposed during the evolution of recommendation systems. These recommendation models are normally classified into two main categories, which are the content-based and collaborative filtering (CF)-based techniques. The CF is known as the

most popular method which is applied in existing recommendation systems.

In general, the CF-based recommendation systems are designed to predict user's utility for unseen items by analysing other users' historical interactors with these given items. There are several collaborative filtering-based frameworks have been proposed at that time. These collaborative filtering-based techniques are commonly applied in many real-world applications, such as the well-known matrix factorization (MF) (e.g. PMF [31], NMF [32], etc.) and factorization machine (FM) [33]. In general, these out-of-the-shelf recommendation techniques have been designed upon a common paradigm of representation learning principle in which the low-ranked representations of all users and items to assist the user's utility process. However, most of the recent CF-based works still suffered major challenges [4, 8] related to user–item interaction data sparsity as well as the cold-start problem. Therefore, many researchers have sought out potential integration of deep learning in current recommendation systems in order to tackle existing challenges. From the bird-eye's views on recent works within the recommendation area which are close to our works in this paper, we categorized these studies into two main categories, which are graph neural network (GNN)-based and deep learning-based representation learning over heterogeneous network (RSHIN)-based recommendation approaches.

2.1 GNN-based recommendation system

In recent years, the graph neural network has been considered as the main stream for most of networked data representation learning. Many GNN-based architectures have supported to achieve state-of-the-art performances in multiple information network analysis and mining tasks. Within the recommendation area, the integration of heterogeneous information networks and GNN-based models for networked data representation learning have been widely attended recently. This integrated network embedding approach supports to produce user and item embeddings which are utilized as the auxiliary side information. These auxiliary side information resources are highly associated with the recommendation-oriented target entities, such as users and items, which have been widely studied. In the last few years, multiple GNN-based architectures [34] (e.g. GCN [15], GraphSage [16], GAT [17], etc.) have emerged and become a popular and powerful graph-structured data representation learning approach. These advanced graph-based deep learning approaches have demonstrated state-of-the-art performances in multiple areas of networked data analysis and mining. Within the recommendation area, the GNN is also adopted in the process of structural representation learning of user–item

interactions. The interaction data between users and items are modelled as graph-based structure. These graph-based structures have showed the significant enhancements in multiple recommendation downstream tasks. Like as the well-known NGCF model [7] of Wang et al. which proposed a novel user and item representation learning technique for handling recommendation task by using GCN over user–item interaction data in form of the bipartite-based graph structure. Similar to this approach, within the LightGCN model [18], He et al. proposed the utilization of a lightweight version of multi-layered GCN architecture. It supports to efficiently learn the user and item representation through the linearly propagation learning procedure of GCN. Considering the GCN-based architecture as a promising direction for user and item representation learning in recommendation, recently several notable GCN-based recommendation methods have been proposed, like as: SMOG-CF [35], LR-GCCF [36], ESRF [37], A2-GCN [38] and IMP-GCN [39]. These models have been designed upon a common principle of high-order graph-structured representation learning over the user–item interaction data in the form of an information network. The global high-ordered structural features of users and items are mainly utilized to assist the recommendation task. However, most of recent GNN-based recommendation methods are designed to mainly target homogeneous user–item interactive relationships. Thus, they might be unable to integrated with associated heterogeneous auxiliary information in forms of network heterogeneity.

2.2 Deep learning-based recommendation over HIN (RSHIN)

Along with notable progresses in deep learning-based recommendation system with GNN, many recent studies have also focused on the integration between deep learning approach with heterogeneous network modelling and representation learning as side information which supports to enrich the recommendation context. HIN has been emerged and considered as an important concept in modelling and analysing complex graph-structured datasets with the interactions of multi-typed nodes/entities via different types of relationships. In early attempts of integrating HIN with recommendation (e.g. HeteRecom [40], HeteCF [41], SemRec [42], etc.), researchers have tried to characterize preferences of users and items by modelling semantic interactive paths (in forms of meta-paths/meta-graphs) and calculating semantic similarity weights between them in the form of a heterogeneous network. These early efforts have demonstrated significant improvements in the performance of recommendation task as well as proved the potentiality of this researching direction.

The HIN-integrated recommendation model as side information However, these meta-path-based similarity-based models for recommendation problem still lacked the capability of flexibly learning the latent feature representations of users and items. Thus, they can be generally applied for dealing with different recommendation tasks. With the tremendous raises of advanced network embedding techniques, there are multiple integrated HIN embedding for recommendation-based methods. These methods have been proposed to deal with this challenge. Such as the well-known works of Shi et al. in the HERec model [20] and Zhao et al. in the HetNERec model [23]. These models utilized the meta-path-based representation learning technique with custom embedding fusion mechanisms. These data embedding fusion mechanisms have supported to produce rich schematic recommendation-aware user and item representations.

The HIN-integrated GNN for recommendation task In recent times, there are also several efforts on integrating the GNN-based representation learning over HIN. These HIN-based proposals have significantly improve the performance of recommendation task, such as: MAGNN [21] and MetaHIN [22]. In the MAGNN model [21], Fu et al. proposed a novel meta-path-based aggregation strategy in the GCN-based representation learning framework to capture the rich heterogeneous structural and semantic information of user–item interaction relationships. Similar to that, Lu et al. proposed a meta-learning approach to simultaneously model and exploit the rich structural representations of users and items within the context of network heterogeneity. Mainly relying on the context of network embedding-based recommendation, in recent time, Wu et al. proposed a novel self-supervised graph embedding approach for handling recommendation, called as: SGL [30]. In the proposed SGL model, Wu et al. utilized a custom self-discrimination feature extraction mechanism to obtain rich structural latent representations of users and items from their historical interaction data. These interaction data resources are in forms of networked data structure. Then, the rich graph-based structural features of users and items are utilized to fine-tune for handling recommendation problem. However, most of existing deep learning/GNN-based techniques for recommendation over both homogeneous and heterogeneous information networks are still considered as unable to efficiently integrating with heterogeneous auxiliary information as transferrable resources which have been previously produced by explicitly encoding the rich heterogeneous contextual within-entity embeddings in a general task-driven manner. Thus, majorly motivated by achievements of recent works, we proposed a novel pre-training heterogeneous GNN-based framework for enriching the contextual

information of user and item nodes. Then, the achieved pre-trained node and item embeddings are applied in an existing recommendation framework to improve the performance of recommendation in both accuracy and scalability aspects.

3 PreHIN4Rec: heterogeneous pre-training schema for recommendation

In this section, we formally present the background concept, task definition for the heterogeneous pre-training schema for recommendation task (in Sect. 3.1) as well as the methodology and detailed implementations of our proposed PreHIN4Rec model (in Sect. 3.2).

3.1 Background concept and task definition

For the recommendation domain, the traditional user–item interaction relationships are normally represented as the user–item feedback matrix, denoted as: $\mathfrak{R} \in \mathbb{R}^{n \times m}$, with: (n) and (m) are the numbers of users (as: $\mathcal{U} = \{u_i\}_{i=1}^n$) and items (as: $\mathcal{I} = \{i_i\}_{i=1}^m$), respectively. Relying on the type and structure of given recommendation dataset, each entry point/cell of the given user–item feedback matrix might be composed as binary-based values. These value indicate the interactive activities like clicking, check-in, like, etc., or realistic values which presents for the interesting/satisfaction level of users upon specific items (e.g. rating score, etc.). In specific user–item interaction is identified as real-valued data, like ratings, and each user–item feedback entry can be denoted as a direct relationship in the form of a tuple, as: $\langle u, i, r \rangle$ and $\mathfrak{R} = \{\langle u, i, r \rangle\}_{i=1}^{|\mathfrak{R}|}$. The (r) presents for the rating score of a specific user (u) upon an item (i).

Recommendation over the heterogeneous network Within the context of recommendation over heterogeneous network (RSHIN), the traditional user–item interactions are supplemented with extra auxiliary information. These auxiliary information resources are tightly associated with the target user and item nodes. For examples, with product items, we have other associated node types like brand, category, tags, etc., or age, gender, occupation, group, etc., for users. To sufficiently model and handle these auxiliary information resources, the HIN analysis approach [9, 10] is considered as a powerful paradigm to be applied in this case. In HIN-based analysis approach for recommendation, the user–item feedback relations and their associated entities are modelled as a graph-based structure, denoted as: $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R})$. The (\mathcal{V}) and (\mathcal{E}) are the sets of network nodes and edges, respectively. There are two mapping functions, $\phi(\cdot)$ and $\psi(\cdot)$ which support to specify the types of network nodes and edges, as: $\phi: \mathcal{V} \rightarrow \mathcal{A}$ and

$\psi : \mathcal{E} \rightarrow \mathcal{R}$. The \mathcal{A} and \mathcal{R} present for the sets of node and edge types, respectively. By modelling recommendation dataset as a HIN, we have: $\mathcal{U} \subset \mathcal{V}$, $\mathcal{I} \subset \mathcal{V}$ and $\mathfrak{R} \subset \mathcal{E}$. The ultimate goal of our heterogeneous GNN-based pre-training model for recommendation task is to achieve the embedding mapping function. This mechanism is denoted as a mapping function: $f_{pretrain}(\cdot)$, as the following (as shown in the Eq. 1):

$$f_{pretrain} : \mathcal{G} \rightarrow \mathcal{U}, \mathcal{I} \tag{1}$$

In this equation, the $\mathcal{U}, \mathcal{U} \in \mathbb{R}^{n \times d}$ and $\mathcal{I}, \mathcal{I} \in \mathbb{R}^{n \times d}$, present for the d-dimensional embedding matrices of user and item nodes, respectively. In the real-world implementation of recommendation systems, user and item nodes are normally associated with different node types as well as unstructured data source such as textual data (e.g. contents of user’s reviews, descriptions of products, etc.).

Pre-training network embedding approach for recommendation task In this approach, our proposed model in forms of pre-training network embedding schema is mainly designed for recommendation task. Therefore, we only focused on achieve the rich structural representations of user (\mathcal{U}) and item (\mathcal{I}) nodes only. Then the general pre-trained user and item embeddings are later utilized to effectively fine-tune in an existing recommendation model (e.g. MF [5], NCF [6], NGCF [7] or LightGCN [18]). The recommendation process is denoted as a mapping function: $f_{rec}(\cdot)$ which supports for handling the users’ utility prediction on their unseen items, denoted as: $\hat{\mathfrak{R}}$. The overall process can be formulated as the following (as shown in Eq. 2):

$$f_{rec} : \mathcal{U}, \mathcal{I} \rightarrow \hat{\mathfrak{R}} \tag{2}$$

As modelling the user–item interaction data as a heterogeneous network, to efficiently preserve the semantic direct–indirect relationships between user and item nodes with other node types, we formulate them as meta-paths [9]. In general, a meta-path [9, 10] denoted as (\mathcal{P}) is normally composed as a (l)-length network node sequence, denoted as: $\mathcal{A}_1 \xrightarrow{\mathcal{R}_1} \mathcal{A}_2 \xrightarrow{\mathcal{R}_2} \dots \xrightarrow{\mathcal{R}_l} \mathcal{A}_{l+1}$. In fact, in HIN analysis and mining area, the different meta-paths convey different specific semantic meanings of interactive relationships between two nodes. In real-world recommendation dataset, the user and item nodes might have different meta-paths which indicate various meanings (as shown in Table 1). To do this, in a given HIN we have $(\mathfrak{P}^{\mathcal{U}}, \mathfrak{P}^{\mathcal{U}} = \{\mathcal{P}_i\}_{i=1}^{|\mathfrak{P}^{\mathcal{U}}|})$ and $(\mathfrak{P}^{\mathcal{I}}, \mathfrak{P}^{\mathcal{I}} = \{\mathcal{P}_i\}_{i=1}^{|\mathfrak{P}^{\mathcal{I}}|})$ present for the set of user and item meta-paths, respectively in a given HIN (\mathcal{G}).

In recent times, there are several well-known meta-path-based representation learning techniques, like: HIN2Vec

[28], Metapath2Vec [29], etc. These models enable to preserve the semantic proximity-ordered information between multi-typed nodes for a specific meta-path. Thus, with different meta-paths of each user and item node type, we can achieve a set of meta-path embedding matrices, denoted as: $\mathcal{X}^{\mathcal{U}} = \{\mathbf{X}_i^{\mathcal{U}}\}_{i=1}^{|\mathfrak{P}^{\mathcal{U}}|}$ and $\mathcal{X}^{\mathcal{I}} = \{\mathbf{X}_i^{\mathcal{I}}\}_{i=1}^{|\mathfrak{P}^{\mathcal{I}}|}$, respectively. In our proposed GNN-based pre-training schema, we combine the multi-rational graph-structured in previous COM-P [25] with the meta-path-based representations. It supports to produce the joint heterogeneous structural and schematic representations of user and items nodes for enhancing the recommendation performance. Table 2 shows list of notations and their corresponding explanations which are commonly used in our paper.

3.2 Heterogeneous GCN-based pre-training schema for recommendation

In this section, we introduce our proposed heterogeneous graph neural pre-training strategy which is a combination between multi-rational and meta-path-based representation learning for preserving the joint structural and schematic embeddings of users and items for enhancing the recommendation performance.

3.2.1 Meta-path-based representation learning approach

Meta-path-based rich structural/semantic representation learning for user/item interactions To efficiently preserve the multiple semantic interactive information between users and items with extra auxiliary information in form of network heterogeneity, we first apply the Metapath2Vec model [29] to obtain various rich path-based semantic embeddings of target user and item nodes. Given a HIN, $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A}, \mathcal{R})$ with a set of meta-path for each target node type, denoted as: $\mathfrak{P} = \{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_{|\mathfrak{P}|}\}$, for each meta-path (\mathcal{P}), we apply the meta-path-based random walk mechanism [29] to achieve a heterogeneous node sequences, these walks are randomly generated following the distributions (as shown in the Eq. 3a), with (w_t) presents for the (t^{th}) node in a given walk. From the generated node sequences for each node, a.k.a. node context, as: (C_u), we apply the heterogeneous skip-gram architecture and negative sampling strategy [29] to achieve the meta-path-based embedding for each (u) node with is corresponding context (C_u) by optimizing the learning objective as shown in Eq. 3b.

Table 1 Examples of common meta-paths in recommendation dataset

Meta-path	Meaning
U[User]-I[Item]-U	Two users have interacted with the same item
U-O[Occupation]-U	Two users have a same job/occupation
U-G[Gender]-U	Two users have the same gender
I[Item]-C[Category]-I	Two items belong to a same category
I-B[Brand]-I	Two items belong to a same brand
U-I-B-I-U	Two users have interacted with items which belong to a same brand

Table 2 List of notations which are used in our paper

Notation	Explanation
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	A heterogeneous network with (\mathcal{V}) and (\mathcal{E}) are the sets of network nodes and edges, respectively
\mathcal{A} and \mathcal{R}	The sets of network node and edge type, respectively
\mathcal{U} and \mathcal{I}	The sets of user and item nodes, respectively
u and i	A user and an item, respectively
\mathfrak{R} and $\hat{\mathfrak{R}}$	The sets of historical and predicted user–item rating/interaction data
\mathcal{P} and \mathfrak{P}	A meta-path and a set of meta-paths, respectively
\mathcal{U} and \mathcal{I}	The general embedding matrices of user and item nodes
\mathbf{X} and \mathcal{X}	A meta-path-based embedding matrix and a set of meta-path-based embedding matrices, respectively

$$\text{Prob}(w_{(t+1)} = x | w_t = v, \mathcal{P}) = \begin{cases} \frac{1}{\text{Neighbour}^{\mathcal{A}_{(t+1)}}(v)}, & e_{v \sim x} \in \epsilon \text{ and } \phi(x) = \mathcal{A}_{(t+1)} \\ 0, & \text{otherwise} \end{cases} \quad (3a)$$

$$\max_{f_{\text{metapath2vec}}} \log p(C_u | f_{\text{metapath2vec}}(u)) \quad (3b)$$

In Eq. 3b, the $f_{\text{metapath2vec}}(\cdot)$, presents for given meta-path-based representation learning strategy as a general mapping function, as: $f_{\text{metapath2vec}}: \mathcal{V} \rightarrow \mathbf{X} \in \mathbb{R}^{|\mathcal{V}| \times d}$. The parameters of this mapping function are learnt during the neural training process to maximize the training objective as shown in the Eq. 3b. Multiple rich semantic/structural representation learning for recommendation. At the end of this process, we can achieve a set of meta-path-based embedding matrices for each target node type, denoted as: $\mathcal{X} = \{\mathbf{X}_i\}_{i=1}^{|\mathcal{P}|}$. By doing this, we achieve each embedding matrix (\mathbf{X}_i) carry different semantic information through interactive paths of users and items with other node types in forms of the given meta-path (\mathcal{P}_i). These meta-path-based embedding matrices are later utilized to enrich the pre-trained representations of users and items. These representations are obtained through the multi-rational GNN-based aggregation and recommendation-aware fine-tuning process which are described in next parts.

3.2.2 Multi-rational graph-based and meta-path-based embedding fusion

Mainly inspired from the previous COM-P model [25], the ultimate goal of applying multi-rational graph embedding strategy in our pre-training approach is to efficiently capture the latent feature representations of multi-typed nodes which are associated with target user and item nodes as side information. To do this, first of all, we constructed a multi-relational graph, denoted as: $\mathcal{G}_{\text{mult}} = (\mathcal{V}, \mathcal{E}_{\text{mult}}, \mathcal{R})$ [25], from a given HIN (\mathcal{G}), with: $\mathcal{E}_{\text{mult}}$ presents for set of multi-relation-based edges. For each target node type, for example user node, the multi-relation-based edge set are defined as:

$$\mathcal{E}_{\text{mult}}^u = \mathcal{E}^u \cup \{(v, u, \mathcal{R}_t^{-1}) | (u, v, \mathcal{R}_t)\} \cup \{(u, v, T) | u \in \mathcal{U}\}$$

, with: (u, v, \mathcal{R}_t) and (u, v, T) are the direct relationship (as: $\mathcal{R}_t \in \mathcal{R}$) between a specific user node type (u) to another (v) node and the self-loop connection of a specific user (u), respectively. In general, the utilization of multi-relational graph-based transformation for a given HIN can support to sufficiently capture the type-varied direct relationships between user and item nodes with other node types, thus preserving the high-order proximity between these target type nodes. Then, for learning the GCN-based target node embedding, we reapplied the propagation learning process of GCN through different layer in which target node (user/item) latent feature representations are aggregated through

multi-relational neighbourhood edges. This propagation learning process through GCN for each (i^{th}) user node (i) can be formulated as the following (as shown in the Eq. 4):

$$\mathcal{H}_{u_i}^{[l-1]} = ReLU \left(\sum_{\langle u_i, u_j, \mathcal{R}_t \rangle \in \mathcal{E}_{multi}^u} W_{\mathcal{R}_t}^{[l-1]} \cdot \gamma(\mathcal{H}_{u_j}^{[l-1]}, \mathcal{R}_t^{[l-1]}) \right) \tag{4}$$

In this equation, $W_{\mathcal{R}_t}^{[l-1]}$ and $\mathcal{R}_t^{[l-1]}$ are the weighting parameter matrix for a specific relation \mathcal{R}_t between two users (u_i) and (u_j) and the relation embedding of (\mathcal{R}_t), respectively. The $\gamma(\cdot, \cdot)$ is the translation-based composition operator which is majorly inspired from the TransE model [44]. After propagating through (k^{th}) layers, we achieve the final representations of target user and item nodes, denoted as: \mathcal{U}^{MR} and \mathcal{I}^{MR} , as the last hidden state of each given GCN-based architecture, as: $\mathcal{U}^{MR} = \mathcal{H}_{u_i}^{[k],u}$ and $\mathcal{I}^{MR} = \mathcal{H}_{u_i}^{[k],i}$.

3.2.3 Multiple embedding fusion and recommendation task-driven fine-tuning

Meta-path-based user/item embedding fusion For the most GNN-based representation learning approach which is applied for graph-structured dataset, multiple node embeddings can be flexibly learnt and fused into the unified embedding spaces which are later used to fine-tune for specific task. The task-driven learning objective of a given GNN-based architecture can be defined in various ways such as graph reconstruction or corresponding node label prediction. In our heterogeneous GNN-based pre-training paradigm, we intend to incorporate the meta-path and multi-relation-based latent representations of user and item nodes to facilitate the recommendation task. In order to force the given GNN-based model to simultaneously preserve the personalized correlational and characteristic features between user and item nodes within multiple embedding spaces which are later used for recommendation, we train our model with the rating/interaction-based prediction learning objective by taking the dot product of joint meta-path-based and multi-relation-based embedding vectors of user and item nodes (as shown in the Eq. 5).

$$\hat{\mathcal{R}} = (\mathcal{U}^{MR} \cdot (W^{fuse,u} \cdot Concat[\mathcal{X}^u] + b^{fuse,u}))^T \cdot (\mathcal{I}^{MR} \cdot (W^{fuse,i} \cdot Concat[\mathcal{X}^i] + b^{fuse,i})) \tag{5}$$

$$\mathcal{L} = - \sum_{\langle u,i,r \rangle \in \mathcal{R}, \langle u,i,r \rangle \in \mathcal{R}} r \cdot \log(\hat{r}) + (1 - r) \cdot \log(1 - \hat{r}) + \lambda \Theta_{PreHIN4Rec}^2 \tag{6}$$

PreHIN4Rec model training and optimization. In Eq. (5), to softly align and fuse the concatenated meta-path-based embeddings of users and items with the multi-relation-based ones, we apply a linear neural fusion mechanism with the corresponding weight and bias parameters, as: (W^{fuse}) and (b^{fuse}). These parameters of the fusion mechanism are jointly optimized with other parameters of our proposed PreHIN4Rec model by applying the Binary Cross-Entropy (BCE) loss strategy (as shown in the Eq. 6). This training and optimization approach is mainly inherited from the previous work [25]. At the end of this pre-training process, we will achieve the final representations of users and items which can be used as the rich structural and schematic side information resource for the continuously deploying and fine-tuning in different recommendation frameworks in order to achieve better recommendation results in both accuracy and scalability performance aspects.

3.3 Model computational complexity analysis

The complexity of our proposed pre-trained network embedding approach for recommendation is mainly relied on the computational efforts of the Metapath2Vec and two composed meta-path-based graph neural network architectures for jointly learning the representations of users and items. For the Metapath2Vec model, following the previous studies [29, 44], the time and space complexity for generating a set of contextual network nodes for a specific target node through a meta-path is identified as: $\mathcal{O}(|\mathcal{V}|)$ and $\mathcal{O}(l, k)$, with (l) being the length of walk and (k) being the number of taken contextual node samples for a specific target node. In the form of meta-path-based network representation learning for recommendation problem through Metapath2Vec model, the overall time and space complexity of this process are approximately: $\mathcal{O}(|\mathcal{V}|^2)$. These times and computational efforts are required to store and compute the change in model’s parameters to maximize the learning objective as the probability of: $\log p(C_u | f_{metapath2vec}(u))$. For the utilization of GCN-based architecture for learning the global structural features of users and items within their networked interaction data, the time and space complexity for each graph neural network-based architecture is about: $\mathcal{O}(ld|\mathcal{E}| + l \cdot d^2|\mathcal{V}|)$ [45]. Within the joint representation learning of Metapath2Vec and GCN as a pre-trained network embedding schema, once the user and item representations are pre-trained, they can be later reused for fine-tuning with out-of-the-shelf recommendation framework to ensure the effectiveness and stability.

4 Experiments and discussions

In this section, we present extensive experiments in the benchmark MovieLens and Foursquare datasets to demonstrate the effectiveness of our PreHIN4Rec model with recent state-of-the-art baselines, including: Metapath2Vec [29], NCF [6], NGCF [7], LightGCN [18] and SGL [30] for handling the top-n recommendation task.

4.1 Datasets and experimental settings

4.1.1 Dataset descriptions

To evaluate the performance of different recommendation techniques, including our PreHIN4Rec model and other comparative methods, we mainly used two datasets, which are:

- *MovieLens*¹: is a well-known movie recommendation dataset which is released by GroupLens. This dataset has multiple versions (e.g. MovieLens-100 K, MovieLens-1 M, etc.) which are different in size. For experiments in this paper, we mainly used the MovieLens-1 M version. Basically, this dataset contains a set of user's ratings on specific movies, it also has information about other associated node types such as age, gender, occupation for users and genre for movies. MovieLens is considered as a heterogeneous network with multi-typed nodes and links which is suitable for evaluating RSHIN-based and heterogeneous network pre-training-based recommendation techniques.
- *Foursquare*²: is a well-known social location/venue dataset which contains the check-in interaction of users in specific locations/venues. The Foursquare dataset is also considered as heterogeneous network with extra auxiliary information associated to target user and item (location/venue) nodes, such as: tags and tips in forms of textual data.
- *Amazon-Book*: is also a well-known e-commercial product recommendation dataset which is provided by McAuley, J., et al. [46] at this online repository³. This dataset is constructed by collecting over 12 M customer's reviews from products of Book-related categories. This dataset is considered as a heterogeneous network with different types of nodes (user, item and product's category) and their relationships. For experiments in this paper, we randomly selected 100 K user's

ratings from this dataset and constructed the corresponding heterogeneous network.

Dataset pre-processing steps Table 3 shows general statistical information of MovieLens and Foursquare datasets. For the setups of datasets in forms of heterogeneous networks, we modelled the MovieLens dataset as a heterogeneous network with 5 different node types, which are: "user", "movie", "occupation", "gender", "age". For the Foursquare dataset, we constructed a HIN with three node types, as: "user", "location" and "tag". For the Amazon-Book dataset, the constructed HIN also contains three types of network's nodes which are: "user", "item" and book product related "category". The sets of meta-paths which are associated with different interactive semantic meanings of user and item nodes are listed in Table 4. These meta-paths are used to learn the meta-path-based representations of user and item nodes through the Metapath2Vec model [29].

4.1.2 Environmental setups and configurations

For the setup of our proposed PreHIN4Rec model, we mainly used the Python programming language with the support of PyTorch machine learning platform to implement the PreHIN4Rec model. For the experimental environment, our model and other comparative baselines (listed in Sect. 4.2) are set up in a single server with the Nvidia Tesla K80 24 GB GDDR5 GPU. For the configurations on GCN-based architecture in our pre-training schema (described in 3.2.2), we set the number of GCN-based layer to 3 ($k^{GCN} = 3$) and general dimensionality of network node embedding vector is set to 64 ($d^{GCN} = 64$). For the implementation of Metapath2Vec model in our approach (described in 3.2.1), we utilized the original implementation of Dong, Y., et al. [29]⁴ with the number of walk per node (n_{wpn}) and walk's length (wl) is set to 5 and 100 ($n_{wpn} = 5, wl = 100$), respectively. The dimensionality of meta-path-based embedding vector is set to 300 ($d^{mp2vec} = 300$). The other experimental configurations of our proposed PreHIN4Rec are specified in Table 5.

Similar to recent works in the GCN-P/COM-P [25], as the heterogeneous GCN-based pre-training approach, we utilized our PreHIN4Rec model to learn the user and item embeddings for recommendation task in a general manner. Then, the pre-trained models are continuously applied as the initial state and fine-tuned for handling top-n recommendation problem with existing recommendation frameworks, including: MF [5], named as: (*PreHIN4Rec*^{MF}), NCF [6], named as: (*PreHIN4Rec*^{NCF}), NGCF [7], named

¹ MovieLens dataset: <https://grouplens.org/datasets/movielens/>.

² Foursquare dataset: <https://sites.google.com/site/yangdingqi/home/foursquare-dataset>.

³ Amazon-Book dataset: <http://snap.stanford.edu/data/web-Amazon-links.html>.

⁴ Metapath2Vec: <https://ericdongyx.github.io/metapath2vec/m2v.html>.

Table 3 General summarization of MovieLens and Foursquare datasets

Dataset	No. users	No. items	No. other node's types	Density	Total		
					Nodes	Edges	
MovieLens-1 M	6040	3883	Gender	2	0.02061619	9971	1,024,737
			Occupation	21			
			Age	7			
			Genre	18			
Foursquare	2060	2876	Tag	4179	0.00041331	9115	35,294
Amazon-Book	64,089	5672	Category	2326	0.00003943	72,087	102,462

Table 4 List of meta-paths used for experiments in MovieLens and Foursquare datasets

Dataset	Target node type	Meta-paths
MovieLens	User	U[User]-I[Item]-U, U-O[Occupation]-U, U-G[Gender]-U, U-A[Age]-U
	Item	I[Item]-Ge[Genre]-I, I-U-I
Foursquare	User	U-I(check-in)-U, U-I(tip)-U
	Item	I-T[Tag]-I, I-U(check-in) ⁻¹ -I, I-U(tip) ⁻¹ -I
Amazon-Book	User	U-I-U, U-I-C[Category]-I-U
	Item	I-U-I, I-C-I

Table 5 List of experimental configurations for PreHIN4Rec model

Parameter	Value
Number of hidden layers for the GCN-based pre-training mechanism (k^{GCN})	3
Dimensionality of GCN-based embedding vector (d^{GCN})	64
Metapath2Vec [29]: number of walks per node (n_{wpm})	5
Metapath2Vec [29]: walk length (wl)	100
Metapath2Vec [29]: dimensionality of embedding vector (d^{mp2vec})	300
Metapath2Vec [29]: number of training epochs for each meta-path ($epoch^{mp2vec}$)	60
Maximum number of training epochs for GCN-based pre-training mechanism ($epoch^{GCN}$)	300
Early stopping after number of epochs	10
General learning rate (η)	1×10^{-4}
Model's optimizer	Adam (weighting decay: 5×10^{-3})

as: ($PreHIN4Rec^{NGCF}$) and LightGCN [18], named as: ($PreHIN4Rec^{LightGCN}$).

Evaluation metric usage As targeting to deal with the top-n recommendation problem, to evaluate the recommendation outputs by using different techniques, we mainly used the normalized Discounted Cumulative Gain (nDCG) and Mean Average Precision (MAP) evaluation metrics.

4.2 Comparative methods

To evaluate the performance of our proposed heterogeneous GCN-based pre-training paradigm for

recommendation task with recent state-of-the-art recommendation baselines, we implemented several techniques for the comparative purposes, which are:

- $MF + BPR$ [5, 47]: is the combination between the traditional matrix factorization framework with the Bayesian personalized ranking (BPR) [47] for handling top-n recommendation task. The BPR-based method applied the generic optimization strategy, called as BPR-Opt to maximize the posterior estimations which are derived from the Bayesian analysis process, thus can be effectively optimized for recommended item ranking problem.

- *Metapath2Vec* [29]: is a classical meta-path-based network representation learning in heterogeneous information network. This model is firstly proposed by Dong, Y. et al. which enables to sufficiently preserve the rich heterogeneous structural features from user–item interactions within the context of network heterogeneity. In order to generate a heterogeneous node sequence for a specific node, the Metapath2Vec utilizes the meta-path-based random walk mechanism. The generated heterogeneous node sequences are later fed into a full-connected neural architecture to learn the node representations within the Skip-gram architecture. For the experiments in this paper, we used the Metapath2Vec to learn the user and item and node embeddings through different meta-paths and then facilitate the MF framework for conducting recommendation with average node embedding vectors of users and items.
- *NCF* [6]⁵: is considered as an early neural network-based recommendation approach for handling n-top-ranking recommendation problem. In the NCF model, He et al. [6] proposed the utilization of neural network-based interpretation mechanism to efficiently learn the personalized user and item embeddings for facilitating the matrix factorization framework. Through experiments, NCF has demonstrated significant improvements in personalized top-n ranking recommendation task. NCF is also considered as a notable baseline for most recent neural collaborative filtering techniques [7, 18].
- *NGCF* [7]⁶: is recently proposed by Wang et al. which is majorly inspired from the user and item collaborative representation learning paradigm of NCF [6]. In the NGCF model, Wang et al. utilized the graph neural network embedding approach to efficiently preserve the interactions between users and items in the bipartite graph-structured form. By utilizing graph neural network-based architecture like GCN, NGCF model [7] can sufficiently capture the high-ordered proximity features between users and items to facilitate the user’s utility prediction process.
- *LightGCN* [18]⁷: is built upon the neural representation learning concept of NCF [6] and NGCF [7], the LightGCN model is considered as a lightweight version of GCN-based architecture for dealing with recommendation task with fewer neural network-based components in which can support to effectively reduce the

computational efforts. In the LightGCN model, He et al. proposed a novel normalized sum of neighbour representation learning strategy for learning the recommendation-aware user and items embeddings which supports to significantly improve both accuracy and time-consuming performances of recommendation outputs in comparison with previous neural collaborative recommendation baselines [6, 7].

- *SGL* [30]⁸: is a recent enhanced graph-based representation learning approach for recommendation problem with the custom self-supervised learning mechanism. In the SGL model, Wu, J. et al. proposed a novel self-supervised network embedding mechanism to effectively explore latent features from user–item interactions to enhance the outputs for personalized recommendation task. In order to do this, the SGL model utilizes the self-discrimination feature extraction mechanism within the reinforcement learning paradigm to obtain the rich structural user and item representations from the historical interactions. Extensive experiments in benchmark datasets demonstrated the outperformance of SGL in comparing with previous graph embedding-based methods in dealing with personalized recommendation problem.

For the configurations of these aforementioned recommendation techniques, we set up them the same as described in the original published papers and for common configurations with our proposed PreHIN4Rec, we configured them as the same with our model which are listed in Table 5.

5 Results and discussion

For the setup of experiments which are presented in this section, we implement several traditional (MF) as well as state-of-the-art deep learning-based recommendation techniques (e.g. Metapath2Vec, NCF, NGCF, LightGCN and SGL). These models are implemented to handle the top-ranking recommendation problem. The experimental outputs of all recommendation techniques are assessed under the MAP (Mean Average Precision) and nDCG (normalized discounted cumulative gain) standard evaluation metrics. For all experiments with each model in different datasets, we run the experimental five times and reported the average result as the final model’s performance output.

The Figs. 2, 3 and 4 show the experimental outputs in terms of nDCG and MAP evaluation metrics with different techniques within the benchmark MovieLens-1 M,

⁵ NCF model: https://github.com/hexiangnan/neural_collaborative_filtering.

⁶ NGCF model: https://github.com/xiangwang1223/neural_graph_collaborative_filtering.

⁷ LightGCN model: <https://github.com/kuandeng/LightGCN> (Tensorflow), <https://github.com/gusye1234/LightGCN-PyTorch> (PyTorch).

⁸ SGL model: <https://github.com/wujcan/SGL>.

Fig. 2 Experimental outputs in terms of nDCG and MAP evaluation metrics for top-n recommendation task with different baselines in the MovieLens-1 M dataset

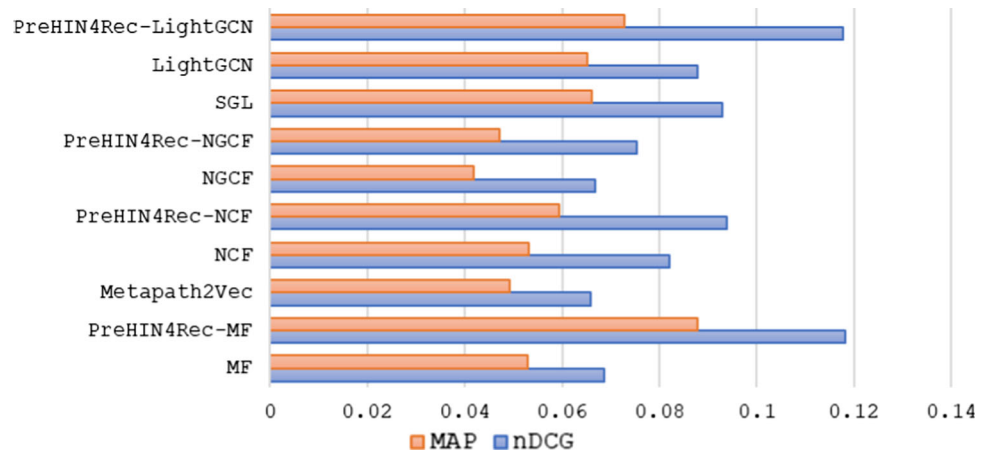


Fig. 3 Experimental outputs in terms of nDCG and MAP evaluation metrics for top-n recommendation task with different baselines in the Foursquare dataset

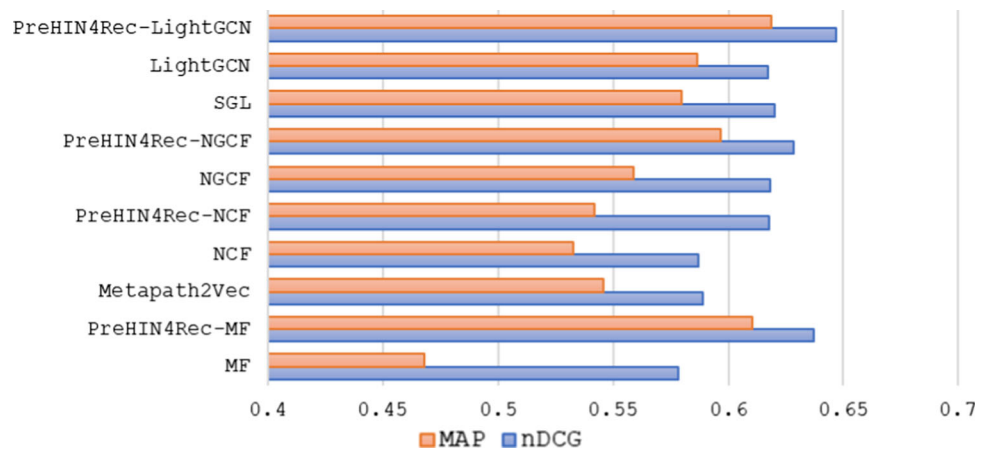
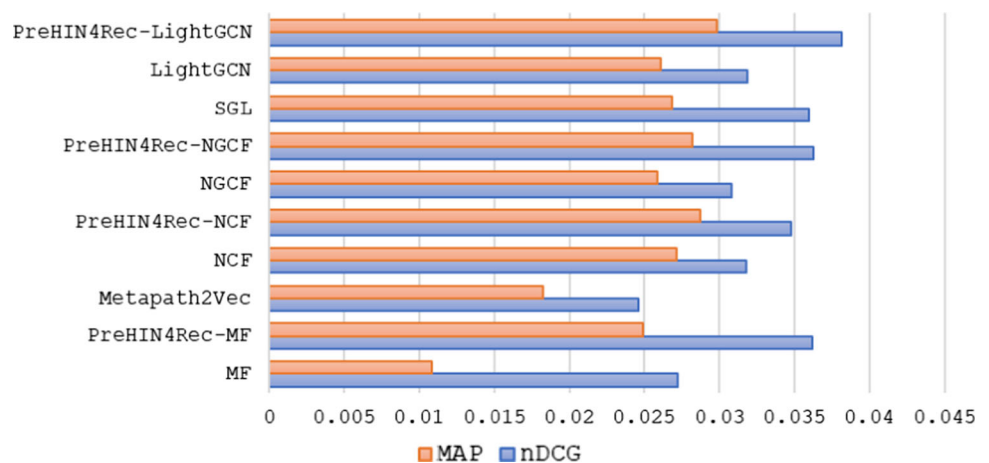


Fig. 4 Experimental outputs in terms of nDCG and MAP evaluation metrics for top-n recommendation task with different baselines in the Amazon-Book dataset



Foursquare and Amazon-Book datasets, respectively. In general, as shown from the experimental outputs, our proposed PreHIN4Rec model can support to achieve better performances as the transferrable knowledge resources for existing recommendation techniques than the original versions. Specifically, in the MovieLens-1 M dataset, with the supports of our PreHIN4Rec pre-training mechanism, the previous recommendation platforms significantly

improved the accuracy performances about: 72%/66.4% (MF), 53.64%/35.73% (Metapath2Vec), 14.2%/11.6% (NCF), 12.8%/12.8% (NGCF), 33.9%/11.7% (LightGCN) and 8.96%/1.02% (SGL) in terms of nDCG and MAP evaluation metrics, respectively. As shown from the experimental results our model significantly improved nearly double the recommendation performances in terms of MAP and nDCG of the traditional MF technique. In

addition, other state-of-the-art network embedding-based techniques also remarkably achieve better performances than the classical MF and thus demonstrate the necessary of both deep learning/transfer learning-based approaches for leveraging the recommendation outputs within the benchmark MovieLens dataset.

Within the Foursquare dataset, similar to experimental results in previous MovieLens dataset (as shown in Fig. 3), our proposed PreHIN4Rec model also supported to slightly enhance the accuracy performances for existing recommendation methods, about: 10.2%/30.4% (MF), 7.35%/8.48% (Metapath2Vec), 5.2%/1.7% (NCF), 1.7%/6.7% (NGCF), 4.8%/5.5% (LightGCN) and 2%/2.1% (SGL) in terms of nDCG and MAP evaluation metrics, respectively. The experimental results also presented the effectiveness of our proposed model in the Amazon-Book dataset (as shown in Fig. 4). In general, our proposed pre-trained network embedding approach achieve better performance about 32.9%/129% (MF), 47.61%/53.25% (Metapath2Vec), 9.4%/5.9% (NCF), 17.7%/8.9% (NGCF), 19.7%/14.2% (LightGCN) and 1.1%/3.94% (SGL) in terms of nDCG and MAP-based accuracy scores. These experimental outputs demonstrated the usefulness and potentiality in the utilization of graph neural pre-training paradigm on the existing neural collaborative filtering approach which can effectively improve the top-n recommendation outputs in both accuracy, stability and scalability aspects. Moreover, in our proposed PreHIN4Rec model, we also use the powers of meta-path-based representation learning mechanism via the Metapath2Vec model to enrich the contextual information of user and item embeddings by integrating with semantic interactive information of paths between users, items and other associated node types in form of network heterogeneity.

To sum up, the experimental results within all benchmark recommendation datasets have demonstrated the outperformance of our proposed PreHIN4Rec model in this paper in comparing with both traditional as well as state-of-the-art deep learning-based techniques. These empirical outputs have proved not only the effectiveness of applying heterogeneous network representation learning context through different meta-paths as side information for recommendation systems but also the necessary of adapting pre-trained heterogeneous network schema for improving the stability as well as accuracy performance of recommendation problem within the context of sparse/large-scaled networked data resources, such as social networks, e-commercial networks, etc.

5.1 Ablation studies

In this section, we presented extensive parameter sensitivity studies on our proposed PreHIN4Rec model. As a

neural pre-training approach, there are a huge number of tuning parameters which are required to be taken in consideration in order to make our integrated recommendation systems can achieve the adequate outputs. In our approach, there are two main components which are the meta-path-based embedding mechanism and the GCN-based architecture to handle the overall graph neural representation learning process for dealing with recommendation task. The multi-layered GCN-based architecture is considered as the most important component in which is in charged for the overall structural and recommendation-aware contextual information preserving. Thus, we first study the influences of number of used GCN-based layers, as: (k^{GCN}) as well as the dimensionality of output user and item embedding vectors, as: (d^{GCN}). To do this, we separately varied the values of these two parameters in the ranges of [1–5] and [10–100] for the (k^{GCN}) and (d^{GCN}), respectively, while fixing the other model's parameters and reported the changes in experimental outputs in terms of nDCG metric for top-n recommendation task. For these experiments, we mainly used the Foursquare dataset. The fluctuations of accuracy outputs in terms of nDCG metric (as shown in Fig. 5) of our proposed PreHIN4Rec pre-training schema on different recommendation techniques showed that our model is quite insensitive with these two parameters. Most of the integrated PreHIN4Rec recommendation techniques, such as: PreHIN4Rec^{MF}, PreHIN4Rec^{NCF}, PreHIN4Rec^{NGCF} and PreHIN4Rec^{LightGCN} in which these techniques achieve the stability with $k^{GCN} \geq 3$ (as shown in Fig. 5a) and $d^{GCN} \geq 60$ (as shown in Fig. 5b).

Similar to that, we also conducted extra empirical studies on the influences of meta-path-based embedding mechanism via Metapath2Vec model upon the quality of user/item embedding vector as well as overall recommendation performances of our PreHIN4Rec model. For the meta-path-based embedding mechanism, we studied the effects of embedding vector size, denoted as: (d^{mp2vec}) and the number of training epochs, denoted as: ($epoch^{mp2vec}$). We varied the values of these two (d^{mp2vec}) and ($epoch^{mp2vec}$) parameters in ranges of [10–400] and [10–100], respectively. Experimental outputs in Foursquare dataset presented that our PreHIN4Rec pre-training schema can support existing recommendation platforms to achieve the convergence points with the value of $d^{mp2vec} \geq 300$ (as shown in Fig. 6a) and $epoch^{mp2vec} \geq 60$. (as shown in Fig. 6b).

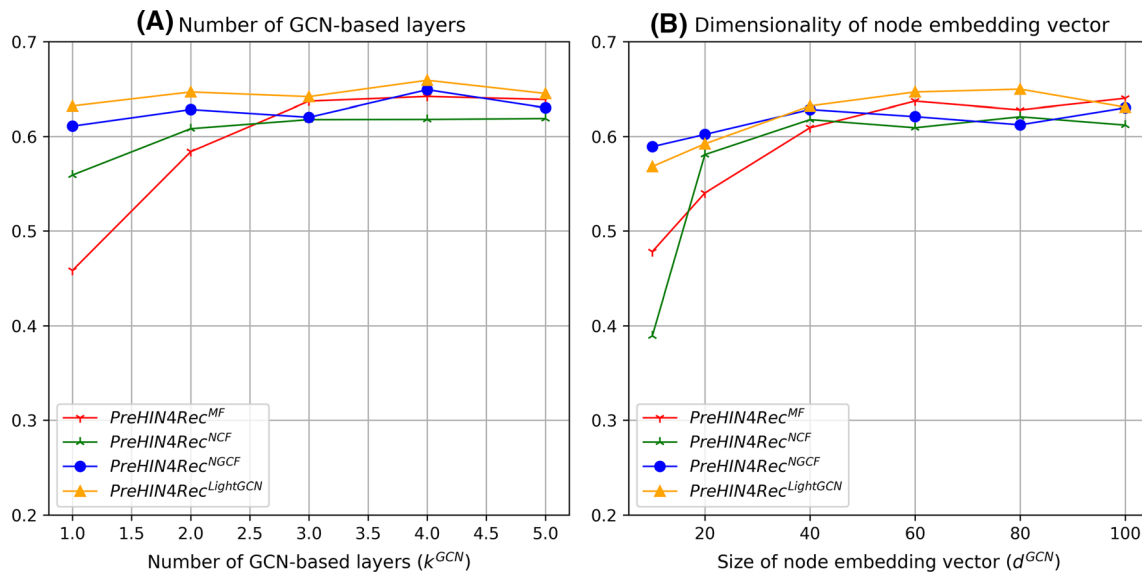


Fig. 5 Model’s parameter sensitivity studies on the number of GCN-based layers (k^{GCN}) and dimensionality of node embedding vector (d^{GCN})

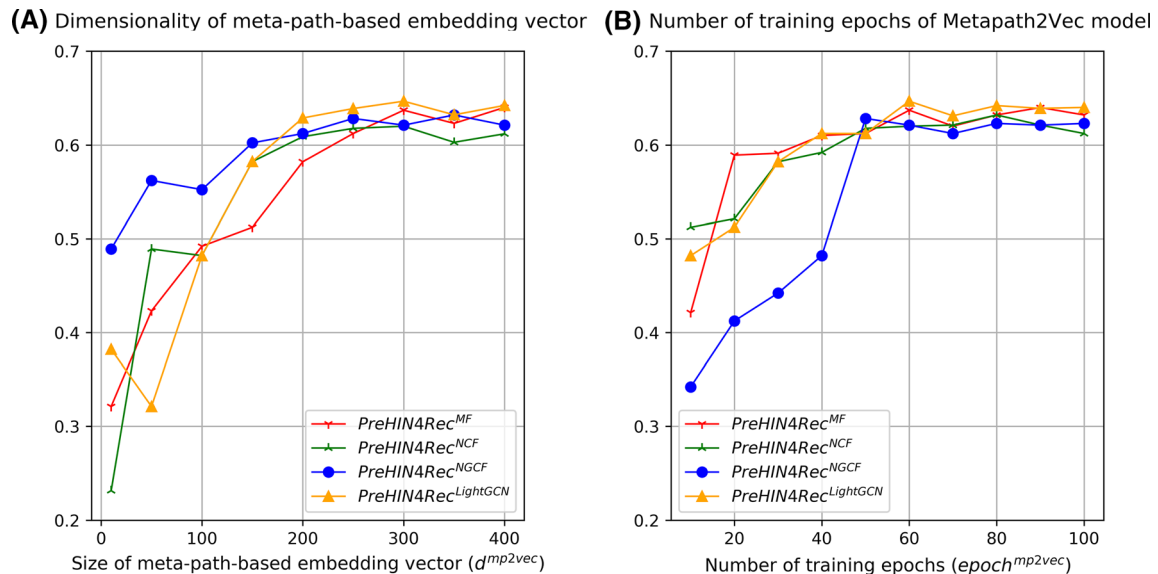


Fig. 6 Model’s parameter sensitivity studies on the dimensionality of meta-path-based embedding vector (d^{mp2vec}) and number of training epochs for Metapath2Vec model ($epoch^{mp2vec}$)

6 Conclusion and future works

In this paper, we presented a novel heterogeneous graph neural pre-training schema for enhancing the performance of recommendation task, called as PreHIN4Rec. In our PreHIN4Rec model, we presented a novel combination between meta-path-based and multi-rational neural graph representation learning approach. It can support to efficiently capture the rich heterogeneous contextual and structural information from user–item interaction data in forms of network heterogeneity. Normally, the existing neural collaborative filtering techniques majorly

encountered challenges. These challenges are mainly related to stability with the variety in modes initialization stages as well as the capability of integrating with pre-trained transferrable knowledge resources to better fine-tune for recommendation task. Therefore, in this paper we proposed the PreHIN4Rec as a recommendation-driven neural pre-training solution which can assist to learn the rich structural representations of users and items in a general manner. These achieved pre-trained user and item embeddings are later fine-tuned with existing recommendation platforms, such as: MF, Metapath2Vec, NCF, NGCF, LightGCN and SGL for effectively dealing with

multiple downstream recommendation tasks. Extensive experiments in benchmark datasets have shown the effectiveness of the integration between our proposed heterogeneous graph neural pre-training schema with existing recommendation platforms in improving both accuracy, stability and scalability performances for recommendation task. In more specific, our proposed PreHIN4Rec model has supported to enhance the nDCG accuracy-based performance of personalized recommendation task averagely 19.4% and 2.35% for all benchmark datasets in comparing with our main competitors in this paper, which are: LightGCN and SGL, respectively. For our future works, we intend to expand the graph neural pre-training schema of the PreHIN4Rec to handle other primitive tasks of network analysis and mining domain, such as: network completion and link prediction. Furthermore, we also intend to improve the capability of PreHIN4Rec model on coping with the network representation learning problem in the context of dynamism.

Funding Viet Nam National University Ho Chi Minh City , DS2020-26-01, Phuc Do.

Declarations

Conflict of interest This statement is to certify that all authors have seen and approved the manuscript being submitted. We warrant that we have no conflict of interest.

References

- Zhao X, Zhang Z, Bi X, Sun Y (2020) A new point-of-interest group recommendation method in location-based social networks. *Neural Comput Appl* 8:1–12
- Tahmasebi H, Ravanmehr R, Mohamadrezaei R (2021) Social movie recommender system based on deep autoencoder network using Twitter data. *Neural Comput Appl* 33(5):1607–1623
- Zhang L (2021) Top-N recommendation algorithm integrated neural network. *Neural Comput Appl* 33(9):3881–3889
- Zhang S, Yao L, Sun A, Tay Y (2019) Deep learning based recommender system: a survey and new perspectives. *ACM Comput Surveys (CSUR)* 52(1):1–38
- Rendle S, Krichene W, Zhang L, Anderson J (2020) Neural collaborative filtering vs. matrix factorization revisited. *RecSys*
- He X, Liao L, Zhang H, Nie L, Hu X, Chua TS (2017) Neural collaborative filtering. *WWW*
- Wang X, He X, Wang M, Feng F, Chua TS (2019) Neural graph collaborative filtering. *SIGIR*
- Zhang Z, Cui P, Zhu W (2020) Deep learning on graphs: a survey. *IEEE Trans Knowl Data Eng*
- Sun Y, Han J, Yan X, Yu PS, Wu T (2011) Pathsim: meta path-based top-k similarity search in heterogeneous information networks. *VLDB*
- Shi C, Li Y, Zhang J, Sun Y, Philip SY (2016) A survey of heterogeneous information network analysis. *IEEE Trans Knowl Data Eng* 29(1):17–37
- Pham P, Do P (2019) W-MetaPath2Vec: The topic-driven meta-path-based model for large-scaled content-based heterogeneous information network representation learning. *Expert Syst Appl* 123:328–344
- Do P, Pham P (2021) W-KG2Vec: a weighted text-enhanced meta-path-based knowledge graph embedding for similarity search. *Neural Comput Appl* 33:1–23
- Pham P, Do P (2021) W-MMP2Vec: topic-driven network embedding model for link prediction in content-based heterogeneous information network. *Intel Data Anal* 25(3):711–738
- Pham P, Do P (2020) W-Com2Vec: a topic-driven meta-path-based intra-community embedding for content-based heterogeneous information network. *Intel Data Anal* 24(5):1207–1233
- Kipf TN, Welling M (2016) Semi-supervised classification with graph convolutional networks. *ICLR*
- Hamilton WL, Ying R, Leskovec J (2017) Inductive representation learning on large graphs. *NIPS*
- Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y (2018) Graph attention networks. *ICLR*
- He X, Deng K, Wang X, Li Y, Zhang Y, Wang M (2020) Lightgcn: Simplifying and powering graph convolution network for recommendation. *SIGIR*
- Zhao H, Yao Q, Li J, Song Y, Lee DL (2017) Meta-graph based recommendation fusion over heterogeneous information networks. *SIGKDD*
- Shi C, Hu B, Zhao WX, Philip SY (2018) Heterogeneous information network embedding for recommendation. *IEEE Trans Knowl Data Eng* 31(2):357–370
- Fu X, Zhang J, Meng Z, King I (2020) MAGNN: metapath aggregated graph neural network for heterogeneous graph embedding. *WWW*
- Lu Y, Fang Y, Shi C (2020) Meta-learning on heterogeneous information networks for cold-start recommendation. *SIGKDD*
- Zhao Z, Zhang X, Zhou H, Li C, Gong M, Wang Y (2020) HetNERec: Heterogeneous network embedding based recommendation. *Knowl-Based Syst* 204:106218
- Pham P, Do P (2020) W-Metagraph2Vec: a novel approval of enriched schematic topic-driven heterogeneous information network embedding. *Int J Mach Learn Cybern* 11(8):1855–1874
- Meng Z, Liu S, Macdonald C, Ounis I (2021) Graph neural pre-training for enhancing recommendations using side information. *arXiv preprint arXiv:2107.03936*.
- Jiang X, Lu Y, Fang Y, Shi C (2021) Contrastive pre-training of GNNs on heterogeneous graphs. In: *Proceedings of the 30th ACM international conference on information knowledge management*
- Yang C, Xiao Y, Zhang Y, Sun Y, Han J (2020) Heterogeneous network representation learning: a unified framework with survey and benchmark. *IEEE Trans Knowl Data Eng*
- Fu T. Y, Lee W. C, Lei Z (2017) Hin2vec: Explore meta-paths in heterogeneous information networks for representation learning. *CIKM*
- Dong Y, Chawla N. V, Swami A (2017) metapath2vec: Scalable representation learning for heterogeneous networks. *SIGKDD*
- Wu J, Wang X, Feng F, He X, Chen L, Lian J, Xie X (2021) Self-supervised graph learning for recommendation. In: *Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval*
- Mnih A, Salakhutdinov RR (2008) Probabilistic matrix factorization. *Advances in neural information processing systems*
- Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755):788–791
- Rendle S (2012) Factorization machines with libfm. *ACM Trans Intel Syst Technol (TIST)* 3(3):1–22

34. Wu Z, Pan S, Chen F, Long G, Zhang C, Philip SY (2020) A comprehensive survey on graph neural networks. *IEEE Trans Neural Networks Learn Syst* 32(1):4–24
35. Zhang H, McAuley J (2020) Stacked mixed-order graph convolutional networks for collaborative filtering. *SDM*
36. Chen L, Wu L, Hong R, Zhang K, Wang M (2020) Revisiting graph based collaborative filtering: a linear residual graph convolutional network approach. *AAAI*
37. Yu J, Yin H, Li J, Gao M, Huang Z, Cui L (2020) Enhance social recommendation with adversarial graph convolutional networks. *IEEE Trans Knowl Data Eng*
38. Liu F, Cheng Z, Zhu L, Liu C, Nie L (2020) A2-GCN: an attribute-aware attentive GCN model for recommendation. *IEEE Trans Knowl Data Eng*
39. Liu F, Cheng Z, Zhu L, Gao Z, Nie L (2021) Interest-aware message-passing GCN for recommendation. In: *Proceedings of the web conference 2021*
40. Shi C, Zhou C, Kong X, Yu P. S, Liu G, Wang B (2012) Heterocom: a semantic-based recommendation system in heterogeneous networks. *SIGKDD*
41. Luo C, Pang W, Wang Z, Lin C (2014) Hete-cf: social-based collaborative filtering recommendation using heterogeneous relations. *ICDM*
42. Shi C, Zhang Z, Luo P, Yu P. S, Yue Y, Wu B (2015) Semantic path based personalized recommendation on weighted heterogeneous information networks. *CIKM*
43. Bordes A, Usunier N, Garcia-Duran A, Weston J, Yakhnenko O (2013) Translating embeddings for modeling multi-relational. *Advances in neural information processing systems*
44. Grover A, Leskovec J (2016) node2vec: scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*
45. Kipf TN, Welling M (2017) Semi-supervised classification with graph convolutional networks. In: *5th international conference on learning representations ICLR*
46. McAuley J, Leskovec J (2013) Hidden factors and hidden topics: understanding rating dimensions with review text. In: *Proceedings of the 7th ACM conference on recommender systems*
47. Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L (2009) BPR: bayesian personalized ranking from implicit feedback. In: *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*

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