Expert Systems With Applications 206 (2022) 117737

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

Categorization of knowledge graph based recommendation methods and benchmark datasets from the perspectives of application scenarios: A comprehensive survey

Nasrullah Khan [a,e,1], Zongmin Ma [a,b,2,*], Aman Ullah [c,3], Kemal Polat [d,4]

[a] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
[b] Collaborative Innovation Centre of Novel Software Technology and Industrialization Nanjing, 210000, China
[c] School of Computer Science and Engineering, Central South University, Changsha 410083, China
[d] Department of Electrical and Electronics Engineering, Bolu Abant Izzet Baysal University, Bolu 14280, Turkey
[e] Department of Computer Science, COMSATS University Islamabad, Vehari Campus, Vehari 61100, Pakistan

ARTICLE INFO

Keywords:
Categorization
Knowledge graph
Side information
Recommendation methods
Benchmark datasets
Knowledge repositories

ABSTRACT

Recommender Systems (RS) are established to deal with the preferences of users to enhance their experience and interest in innumerable online applications by streamlining the stress persuaded by the reception of excessive information through the recommendation methods. Although researches have put a lot of efforts in making recommendation processes accurate, specific, and personalized; different issues like cold start, data sparsity or gray sheep etc., still pop up in one or the other form of challenges. Recently, exploitation of *Knowledge Graph* (KG)-based data as *Side Information* in recommendation methods has revealed as a sign of resolution to the corresponding challenges; and thus, acquired incredible focus, applicability, and popularity. The incorporation of KG in recommendation has not only effectively alleviated the contrasting challenges, but also has provided specific, accurate, personalized and explainable recommendations about the target items to the end users. In this paper, we explore well-known RSs, popular knowledge repositories, benchmark datasets, recommendation methods, and future research dimensions about the current research. Intuitively, we investigate recommendation methods and associated datasets with respect to the corresponding application scenarios in a categorical way.

1. Introduction

With the current expansion of big data and paced augmentation of internet technology, the storage capacity has exponentially increased with respect to the volume of online dispersed data by providing enormous aid to the information overload. The *"information overload"*, as a problem, is defined by the Business-Dictionary[5] as *"The stress persuaded by reception of excessive amount of information to make a decision and dealing with this information without knowing about the validity of its timespan"*. With the rapid advancement of the internet technology, the products of daily life requirements of human being are largely hooked up by the vendors via social media and e-commerce networks for convenient business and advertisement. Therefore, redundant bulks of data are dragging countless difficulties towards the online activities of the concerned people. Moreover, the transmission speed of information to user is much higher than the reverse speed of interactions, results in drop-ratio of utilization to selection rate due to the information overload; that becomes difficult for individuals to timely select the required options among huge and identical varieties of selectable choices. To feasibly alleviate the crests of uncertainties in selection, it is deemed to streamline the online information overload. Search-engines are exploited to submissively filter out the contents of user interests via specific keywords based on their previous interaction records. This technique is popular and widely applied but the suggestions are specific to the

* Corresponding author at: College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China.
E-mail addresses: nasrullah@nuaa.edu.cn (N. Khan), zongminma@nuaa.edu.cn (Z. Ma), dr.aman@csu.edu.cn (A. Ullah).
1 ORCID: 0000-0002-4942-9583.
2 ORCID: 0000-0001-7780-6473.
3 ORCID: 0000-0002-3999-4917.
4 ORCID: 0000-0003-1840-9958.
5 https://www.businessdictionary.com.

https://doi.org/10.1016/j.eswa.2022.117737
Received 12 August 2021; Received in revised form 30 May 2022; Accepted 31 May 2022
Available online 4 June 2022
0957-4174/

acquired keywords. It cannot effectively generalize the framework to recommend about the potential interests of the users (Shao et al., 2021).

In 1990 s, online assistants – application programs used to filter out irrelevant data – are developed to recommend selectable choices to the end users based on their previous interactions' record. Later on, these application programs are designated as *Recommender Systems* (RSs). The significance of RSs is extraordinarily high due to their utilization in various real-life application scenarios. For instance, more than 35% of the revenue of amazon depends on its own RS (Lee & Hosanagar, 2014). Currently, RSs acquire *Side Information* from KG data-repositories to provide recommendations; and therefore, known as KG-based RSs (Chen et al., 2020b; Li et al., 2021a).

KG stores data in structured and machine readable form by offering easy and simple ways to automatically search, gather, validate and utilize the concerned information. It supports pragmatic approaches to hold and present the bulks of heterogeneous information in triplets – RDF[6] format (Gomez-Perez et al., 2017; Ehrlinger & Wöß, 2016), to feasibly exploit it through experimental processes. Likewise, it effectively plays significant role in the provision of uni-and-multi domain structured information to the corresponding systems. By virtue of KG's vigorous hierarchy, its nodes appear in multiple triplets simultaneously and its edges maintain higher-level relations among entities and entities-to-features as well. Moreover, KG offers powerful path-tracing mechanism to retrieve potential information from the external knowledge repositories. It provides the facility of mapping local entities and relations to the relevant information instances in external KGs to enrich the representations of underlying information (Zhang et al., 2016a; Zhang et al., 2018c).

Scenario-aware recommendation is currently an emerging research-trend in KG-based RSs. There are innumerable application frameworks that exploit KG-based RSs to automatically provide suggestive responses to the end users. For instance, acquiring recommendations about Books or Research-Articles (Yang et al., 2020; Song et al., 2019a; Wang et al., 2019b; Wang et al., 2019f), e-commerce (Ding et al., 2021; Fu et al., 2020; Ma et al., 2019; Zhang & Chen, 2018), News (Liu et al., 2019; Wang et al., 2019b; Wang et al., 2019d; Zhang et al., 2018a), Question-answering interaction (Zhang et al., 2020a; Park et al., 2021; Zheng & Zhang, 2019; Qin et al., 2019), Entertainment[7] (Palumbo et al., 2020; Song et al., 2019a; Hu et al., 2018), MoPI[8] (Burke, 2022; Wang et al., 2020b; Mauro et al., 2020; Dadoun et al., 2019; Sha et al., 2019; Sun et al., 2018b) and Social Connections (Song et al., 2019b; Wang et al., 2018a), etc., are clearly expressive application scenarios.

Although extensive research work is carried out on types, evaluation methods and algorithms of RS; there is lack of an inclusive study on categorization of KG-based recommendation-methods from technical perspectives of utilization in a moderate and comprehensive way. The majority of existing review works explore KG-enhanced RS based on one or the other aspect of utilization and discard its comprehensive categorization with respect to diverse facets of exploitation. For example, (Chen et al., 2016) presented a review on visual-interactive model of the interactive RS that aggregated the visualization technology with RS to well elaborate the explainability of the recommendation. Similarly, (Chen et al., 2018) and (Batmaz et al., 2019) explored the exploitation of CF[9]-algorithms and deep learning models in RS, respectively. Moreover, (Guo et al., 2020) described a broader classification of recommendation methods based on path based methods, embedding based methods and unification based methods. The existing surveys lack categorical representations of recommendations approaches with respect to the underlying implementation techniques, frameworks, datasets' categorization and evaluation protocols. Moreover, the present surveys overlooked the concerns of exploring external knowledge repositories in light of KG-based recommendation approaches.

The objective of this paper is to categorize current and previous research work on KG-based recommendation with respect to the implementation techniques, methods, datasets, frameworks and applied approaches to streamline the available knowledge and clear the future research dimensions. In this work, we analyze and compare literature of about ten years (i.e., 2011–2020) related to KG-embedding-based methods, KG-path-based methods, and their hybrid frameworks (i.e., unification of KG-embedding and path-based methods) exploited to provide recommendations. Moreover, we present a comprehensive up-to-date overview of the-state-of-the-art approaches that applied KG-based recommendation as their core research domain in their works. Our in-depth investigation of KG-based RSs ascertain various potential and significant aspects of the field. For instance, we present domain's preliminaries and background, an indicative touch of explanation-based recommendation methods, knowledge repositories, mapping information to knowledge repositories, KG-exploitation at information level, individual recommendation methods and their combined effects, fairness-aware recommendation and future research dimensions.

The main contributions of this work are enlisted as follows:

- We present a comparative analysis of popular knowledge repositories that are being exploited by the KG-based recommendation methods and the benchmark datasets.
- We categorize and discuss the well-known datasets utilized by KG-based recommendation methods. For easier access, we tabularize the datasets with respect to application scenarios and utilization in year-based publications.
- We categorize past ten year's relevant research work on KG-based recommendation methods with respect to the applied techniques and frameworks of implementation. We tabularize the categories based on the frameworks, the applied datasets, the evaluation protocols and other relevant details like methods, publisher, year, etc.
- We discuss several possible research dimensions to highlight future research-work in KG-based Recommendation methods and their amalgamation with other relevant domains.

The rest of the survey is organized as follows: Section 2 describes the background of KG-based recommendation. Section 3 contains the categorization and details of Benchmark datasets. Section 4 provides the Methodology (i.e., the recommendation techniques used by KG-based RSs). Section 5 delivers future research dimensions; and finally, Section 6 concludes the paper.

2. Background

Prior to a thorough dive in the literature to investigate the recommendation datasets and methods in details; in this section, we define *Knowledge Graph*, describe *KG-based RSs*, and overview relevant external *Knowledge Repositories (KReps)* that are commonly used to enrich KG-based datasets.

2.1. Knowledge graph

It was early 1980 s when the conceptual theory of graph was presented and integrated in the framework of expert systems to enhance the progress of social and medical sciences (Nurdiati & Hoede, 2008). Later on, the provision of organized information of international dates to the intelligent systems was an interesting achievement in the domain (Nilsson, 1996). During the current and preceding decades, the arrival of open linked-data foundation (Bizer et al., 2011) and the representation of heterogeneous knowledge via graph acquired significant attention. Graph-based knowledge is enhanced and widely utilized in logical reasoning, NLP linguistic-domains and semantic web. In 2012, Google coined the term "KG" to entitle the phenomenon and incorporated

[6] Resource Description Framework.

[7] Only contains "Movies" and "Music" in this correspondence.

[8] Matters of Particular Interest.

[9] Collaborative Filtering.

graph-based knowledge into its search engine to effectively map user's queries to the demanded results (Steiner et al., 2012; Ehrlinger & Wöß, 2016; Wang et al., 2018c). Currently, KGs are cured and exploited in different daily-life application-scenarios like recommender systems, search engines, relation detection, and question answering system, etc. (Huang et al., 2019b; Zhang et al., 2020a; Hakkani-Tur et al., 2014). For instance, web-based KGs are capable to provide organized-data[10] to the online information management systems that need fast access to the structured knowledge to avoid unnecessary preprocessing.

KG is a pragmatic information-base that represents bulks of heterogeneous information instances (Ehrlinger & Wöß, 2016) in RDF[11] standard format (Gomez-Perez et al., 2017). RDF and graph data representations are somehow inter-identical where nodes (vertices) are entities and paths (edges) among nodes are relations. RDF triplet format or graph fact, i.e., $< e_h, r, e_t >$, is used to represent a relation r between head-entity e_h and tail-entity e_t. Due to the flexible nature of KG, nodes represent entities that join multiple triplets while edges reflect higher-order relations among entities and hold various number of features of each entity. KG offers intractable paradigm of *path-following* technique to specifically chase, locate and retrieve the required information instances; since KG-based path following procedure is someway similar to the working mechanism of path tracking systems (Dogan et al., 2020).

2.2. KG-based Recommender systems

Collaborative Filtering (CF)-based RSs assume that "People liked what in the past will also like that or same in the future" or "People had a taste in the past will also have that taste in the future" or "two users are as similar as their common activities are similar". CF-based methods are widely applied because these are good in capturing user's preferences and easy to deploy but suffer from the issues of *data-sparsity* – data disperse-ness, *cold-start* – facing the arrival of new or unknown entity, and *gray-sheep* – one who's taste doesn't match with any other's in the scenario (Sun et al., 2019; Cano & Morisio, 2017). Content-based RSs adopt "People liked what with which features in the past, will also like that with those features in the future" as their assumption. In content-based RSs, items are recommended on the similarity of their features with users' interest; but feature-mining degrades the performance. These systems suffer from cold start limitation and abundance of explicit feature extraction (Sun et al., 2019). Moreover, aggregation of interaction and content level similarities is achieved by unification of collaborative and content based RSs (Zhao et al., 2016; Kouki et al., 2019) as hybrid RSs. Although the aggregation of RSs is a possible solution to the mentioned limitations, it results in a trade-off between the complexity and performance. *These RSs are termed as KG-based RSs when they acquire side information from KGs to generate recommendations*.

Currently, *Knowledge Graph* is considered as a strong provider of side information to RSs. To better demonstrate the working mechanism of KG-based RSs, we present an exemplary recommendation scenario in Fig. 1; where TV dramas *"Laal Ishq"* and *"Landa Bazar"* are suggested to both *Ahmad* and *Sana* in-spite of the fact that they are not directly interlinked. Typically, the dramas *"Laal Ishq"* and *"Landa Bazar"* are recommended to *Ahmad* because their writer is *KR Qamar* who's written *"Love, Life and Lahore"* has been watched by him. Similarly, both dramas are recommended to *Sana* because both are starred by *Babar Ali* who's starred *"Wafa"* has been previously watched by her. Therefore, we can conclude that the suggestions disclose clear and explainable inferences of the recommendation process.

2.3. Knowledge repositories

Knowledge repositories are information bases being exploited by subgraphs to enrich local entity representations by mapping their entity information to the relevant information instances in the external knowledge repositories. The scope of the available KReps can generally be divided into two types i.e., *multi* and *uni*-domain-oriented Information Bases (IBs/KGs) based on the type of the knowledge they possess. Multi-domain IBs are general purpose KReps whereas uni-domain IBs contain domain-specific knowledge. For instance, DBpedia (Lehmann et al., 2015), YAGO (Fabian et al., 2007), Freebase (Bollacker et al., 2008), Google's KG (Steiner et al., 2012; Ehrlinger & Wöß, 2016; Wang et al., 2018c), Wikidata (Vrandečić & Krötzsch, 2014), NELL (Carlson et al., 2010), Google's Knowledge-Vault (Dong et al., 2014), PROPSPERA (Nakashole et al., 2011) and DeepDive (De Sa et al., 2016; Zhang et al., 2017) etc. are multi-domain KGs. Similarly, Cyc and Open Cyc, Yahoo's KG, and Facebook's Entity Graph (Sun & Iyer, 2013; Paulheim, 2017), etc. are some more examples of well-known multi-domain IBs. Moreover, KnowLife (Ernst et al., 2014), DSKGMS[12] (Yuan et al., 2020), AKMiner (Huang & Wan, 2013), Con2KG (Goyal et al., 2019) and Bio2RDF (Belleau et al., 2008) are some popular examples of uni-domain KGs. Table 1 summarizes the details of various KReps and KGs that are commonly being utilized in KG-based recommendation systems. Moreover, in this part, we briefly introduce some of the very important external knowledge repositories.

2.3.1. Dbpedia

It was released by Leipzig University in 2007–01 and its data repository is annually updated. It extracts structured information from Wikipedia and creates its KGs. DBpedia's data is inter-linked as key-value pairs and provide openly available large-scaled KG structure on the web. It can be interacted via the standard web browsers, crawlers or query languages like SPARQL, etc. According to the latest release[13] of DBpedia, it contains 13 billion RDF triples where 1.7 billion belongs to English language. English version of DBpedia contains 6.6 million entities where 5.5 million are categorized in the form of consistent ontologies. These continuously updating KReps contain information about 1.5 M persons, 496 K distinct works, 840 K locative places, 286 K firms, 306 K species, 6 K diseases, and 58 K plants, etc. (Lehmann et al., 2015; Paulheim, 2017; Kontokostas et al., 2012; Morsey et al., 2012).

2.3.2. YAGo

YAGO – stands for *Yet Another Great Ontology* – is a product of Saarbrücken Max-Planck-Institute; initially launched in 2008 and its first stable version (i.e., YAGO1) was released by the same institute in 2017–06. YAGO1 automatically extracts information from Wikipedia and other sources like DBpedia (Paulheim, 2017), WordNet (Miller, 1998; Kilgarriff, 2000), GeoNames (Brauner et al., 2007), etc., using various heuristic techniques (Mahdisoltani et al., 2013) and unifies the extracted information in RDF triplets. Its next version (i.e., YAGO2) (Hoffart et al., 2011), relies on curation techniques (Dong et al., 2014) besides information extraction; for example, alignment of GeoNames with GeoWordNet (Giunchiglia et al., 2010) is actually based on manual/automatic curation of entities (Hoffart et al., 2011). Moreover, the 2020-released latest version of YAGO (i.e., YAGO4) dependents on SHACL's semantic constraints to keep the data clean. It contains more than 50 M entities and 2B triples about the stored information (Tanon et al., 2020).

2.3.3. Freebase

It was a publically available and editable Knowledge Base (KB) launched by Metaweb Technologies in 2007 and owned by Google in 2010. About five years later, Google permanently closed it on March 31, 2015 and transferred its data to Wikidata (Paulheim, 2017) which is currently stored on Google's data dumps (Chah, 2017). On 31st of March

[10] The data can be single or multi domain oriented.
[11] Resource Description Framework.
[12] Domain Specific Knowledge Graph with Minimum Supervision.
[13] October 2016.

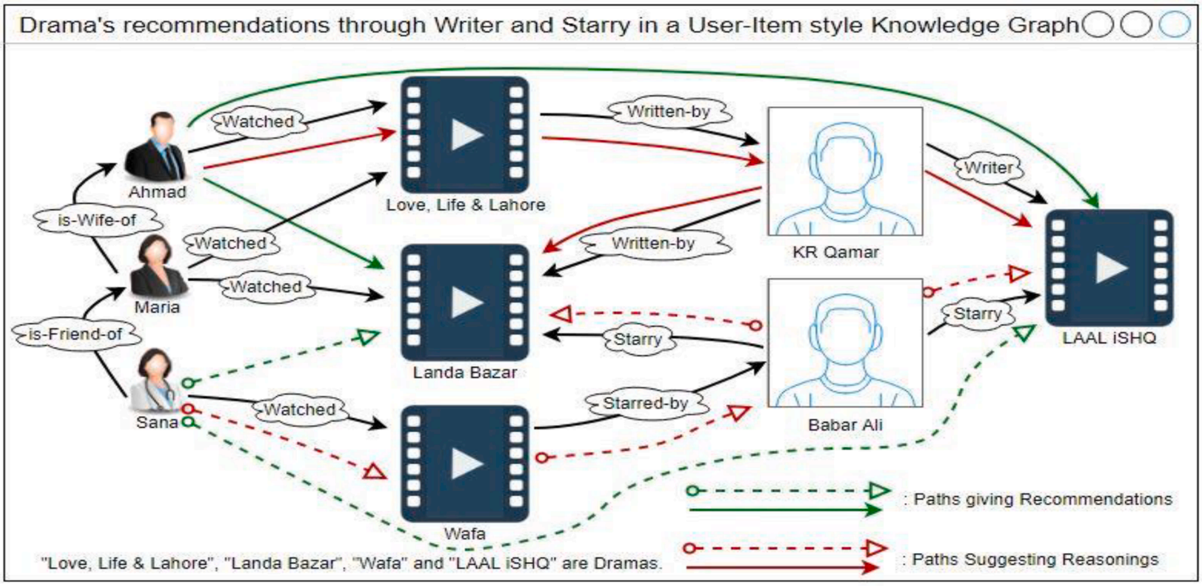

Fig. 1. A disclosure of RSs providing recommendations based on the side information acquired from KG.

Table 1
A comprehensive overview of commonly used KReps and their types, recommendation domains, information source & scope, methods of information collection and references. Abbreviations used – MD: Multi Domain, UD: Uni Domain, IS: Information Scope, KB: Knowledge Base, KS: Knowledge Space, KI: KGs Inventory, KV: Knowledge Vault, EG: Entity Graph, WP: Wikipedia, WN: Word Net, GN: Geo Names, Wd: Wikidata, WD: Web Data, BB: Baidu Baike, HB: Hudong Baike, CW: Chinese Wikipedia, NNDB: Notable Names Database[51], FMD: Flicker Material Database[52], FB: Freebase, EVCF: Edited Version of Curated Freebase KG, MD: Micro-Data[53], GSN: Google's Social Network (Google +), ICLVD: Initially Curated but Latest Versions used Different strategies, FBD: Facebook Data, MB: Music-Brainz[54], ICLVE: Initially Curated but Latest Versions used Extraction as well, PBIDB: Public Bio-Informatics Databases, NCBIDB: National Center for Bio-Technology Information Databases, SL: Scientific Literature, WIP: Web Information Portals, CKB: Cyc Knowledge Base, RVCC: Reduced Version of Curated Cyc-KG, ND: Neurological Disorders, AKB: AETIONOMY's Knowledge Base, GP: General Purpose, HLS: Health and Life Sciences, BI: Bio-Informatics (Biological), CIDBD: Created by Integration of publically available Data from popular Bioinformatics Databases, BM: Bio-Medical, EwMS: Extracted with Minimum Supervision, UBMWB: Unstructured Bio-Medical Web Data, AKMiner: Academic Knowledge Miner, AK: Academic Knowledge, AL: Academic Literature, JB: Jobs Postings, SUHWD: Structured and Unstructured Heterogeneous Web Data.

KReps		Domain		Information			References
Names	Type	MD	UD	Source	IS	Method	
DBpedia	KG	✓		WP	GP	Extracted	(Lehmann et al., 2015)
CN-DBpedia	KG	✓		BB, HB, CW	GP	Extracted	(Xu et al., 2017)
Freebase	KG	✓		WP, NNDB, FMD, MB	GP	Curated	(Bollacker et al., 2008)
YAGO	KG	✓		WP, WN, GN	GP	ICLVE	(Fabian et al., 2007)
Satori	KG	✓		WD	GP	Curated	(Qian, 2013)
Wikidata	KG	✓		WP, FB	GP	EVCF	(Vrandečić & Krötzsch, 2014)
NELL	KG, KB	✓		WD	GP	Extracted	(Carlson et al., 2010)
Google's KG	KG	✓		WD, WP, MD, GSN	GP	ICLVD	(Steiner et al., 2012)
Google's KV	KG, KB	✓		Freebase, WD	GP	Extracted	(Dong et al., 2014)
Yahoo!'s KG	KG	✓		WP, Wd, WD, FB	GP	Extracted	(Blanco et al., 2013)
Facebook's EG	KG	✓		WP, FBD	GP	Extracted	(Sun & Iyer, 2013)
Open Cyc	KG, KB	✓		CKB	GP	RVCC	(Lenat, 1995)
POSPERA	KB	✓		WD	GP	Extracted	(Nakashole et al., 2011)
DeepDive	KB	✓		WD	GP	Extracted	(Niu et al., 2012)
KnowLife	KG		✓	SL, WIP	HLS	Extracted	(Ernst et al., 2014)
DSKGMS	KG		✓	UBMWB	BM	EwMS	(Yuan et al., 2020)
AKMiner	KG		✓	AL, WIP	AK	Extracted	(Huang & Wan, 2013)
Con2KG	KG		✓	SUHWD	JB	Extracted	(Goyal et al., 2019)
Bio2RDF	KS		✓	PBIDB, NCBIDB	BI	CIDBD	(Belleau et al., 2008)
NeuroMMSig	KI		✓	AKB	ND	Curated	(Hoyt et al., 2019)

[51] https://www.nndb.com/.
[52] https://www.fashionmodeldirectory.com/.
[53] https://mdat.com/.
[54] https://musicbrainz.org/.

2015, Freebase contained about 3B triples and 50 M entities (Pellissier Tanon et al., 2016) where more than half of its data was about media (Färber & Rettinger, 2018).

2.3.4. Nell

NELL – Never Ending Language Learning System – is developed by Carnegie Mellon University in 2010. It is currently functional and gradually extends its knowledge base by extracting data[14] from online data-sources via crawlers (Carlson et al., 2010). NELL's data is easily convertible to RDF and accessible in Linked Open Data format (Zimmermann et al., 2013); since it's reasoning capability alleviates the issues of data inconsistency. In the recent version of NELL (i.e., iteration 1002), it contained 2,810,379 instances of different entities with about 1200 categories and 50 K relations among entities. Its ontological structure defines more than 300 classes and about 500 categorical relations among them (Mitchell et al., 2018).

2.3.5. CN-DBpedia

It is released by Fudan University in 2015–12; which is the largest, largely-scaled, multi-domain, structured knowledge repository of Chinese language. It automatically extracts required information in plain-text from Chinese web encyclopedias like Baidu baike, Chinese Wikipedia, Interactive baike, Hudong baike etc.; and filters, unifies and transforms it to the well-structured formats of readability of humans and machines (Xu et al., 2017). Due to its continuous updating nature, frequent variations take place in its entities, facts and rest of the statistics. For instance, in 2019–04, CN-DBpedia2 – the latest version of CN-DBpedia – contained 16,024,656 entities and 228,499,155 triples whereas CN-DBpedia contained only 10,341,196 entities and 88,454,264 triples of information (Xu et al., 2019).

2.3.6. Google's KG

Google's KG – a multi-domain general purpose KG, publically available in 2012 – is the first that introduced the terminology of "Knowledge Graph" to the world of data science. Yet, Google disclosed nothing about how their KG is collected, structured or composed; but a few external resources[15] guessed and expressed – just on the basis of their own experience – that the initial organizational mechanism of Google's KG was curation and later on it adopted the rest of the KG construction and saturation methods wherever required. The resources disclosed that some structured and semi-structured KReps like structured markup Microdata (Meusel et al., 2014) of Schema[16], popular semi-structured web-data such as Wikidata, Wikipedia, and the contents of Google+ (i.e., Google's social media platform), contribute to Google's KG. Google's KG is comprised of 18B statements about 570 M entities; whereas it's schema contains 35 K relation-types that belong to (i.e., connect) 1.5 K entity-types (Dragisic et al., 2014).

2.3.7. Microsoft Satori

MS Satori, launched by Microsoft in 2010–10, is a full-fledged multi-domain general-purpose KG. Although data representation format of satori is RDF, it's publically available and applicable data in RDF-style is limited. MS-Satori is utilized to empower Bing search engine; whose results are therefore, not well comprehensive and generalized (Qian, 2013) as compared to Google's search engine's ranking response. In 2012, MS Satori contained 300 M entities and 800 M relations among them. Currently, (Noy et al., 2019) claims that the knowledge base of MS Satori contains 2B entities and 55B relations among them. Moreover, it has other intermittent KReps like MS Satori standard knowledge base (Szekely et al., 2015), satori3 (Liu et al., 2019), Microsoft Academic Graph (MAG) and Microsoft Academic Knowledge Graph (MAKG) with more than 8B facts of scholarly linked data (Färber, 2019), etc.

3. Benchmark Datasets – Categorization, Embedding and Mapping

KG-based recommender systems are widely applicable in different application scenarios with matchless benefits of interpretability, compatibility, ease of use and accuracy. These method's successful

Table 2
Categorization of application scenarios and utilized data repositories with respect to the references. The *benchmark datasets* are highlighted through the italic text presentation form.

Scenario	Benchmark		Reference(s)
Readings	Douban-Book		(Shi et al., 2018)
	Book-Crossing		(Yang et al., 2020; Tang et al., 2019; Li et al., 2019c; Wang et al., 2019b; Wang et al., 2019c; Wang et al., 2019d; Wang et al., 2019e; Qu et al., 2019)
	Amazon-Book		(Wang et al., 2020b; Qu et al., 2019; Wang et al., 2019f; Huang et al., 2018)
	Intent-Books		(Zhang et al., 2016a)
	DB-books2014		(Cao et al., 2019); (Song et al., 2019a); (Piao & Breslin, 2018)
	LibraryThing		(Palumbo et al., 2020)
	CNRec		(Joseph & Jiang, 2019)
	Bing-News		(Wang et al., 2019b; Wang et al., 2019d, Wang et al., 2018b)
Shopping	*Amazon*		Fu et al. (2020); Ma et al. (2019); Xian et al. (2019); Zhao et al. (2019b); Zhang et al. (2018c), Ai et al. (2018); Zhao et al. (2017)
	Taobao		Ye et al. (2019); Zhao et al. (2019b)
Entertainment	*Movie-Lens-*	100 K	(Xin et al., 2019; Hu et al., 2018; Catherine & Cohen, 2016; Yu et al., 2014; Yu et al., 2013a; Yu et al., 2013b)
			(Palumbo et al., 2020; Yang et al., 2020; Sha et al., 2019; Song et al., 2019a; Cao et al., 2019),
		1 M	(Tang et al., 2019; Qu et al., 2019; Li et al., 2019c; Wang et al., 2019b; Wang et al., 2019d; Wang et al., 2019g; Piao & Breslin, 2018),
			(Huang et al., 2018; Sun et al., 2018b; Palumbo et al., 2018; Palumbo et al., 2017; Zhang et al., 2016a)
		10 M	Nguyen et al. (2020); Tousch (2019); Deldjoo et al. (2018); Hsieh et al. (2017), Ivarsson & Lindgren (2016); Zheng et al. (2016); Dooms et al. (2016)
		20 M	(Huang et al., 2019b, 2018; Qu et al., 2019; Wang et al., 2019b; Wang et al., 2019c; Wang et al., 2019e; Deldjoo et al., 2018)
		25 M	(Penha & Hauff, 2020; Yadav & Vishwakarma, 2020; Leung et al., 2020a; Leung et al., 2020b)
	Douban-Movie		(Shi et al., 2015; Shi et al., 2018; Yang et al., 2020)
	Last-FM		(Palumbo et al., 2020; Yang et al., 2020; Wang et al., 2020b; Song et al., 2019a),(Wang et al., 2019c; Wang et al., 2019d; Wang et al., 2019e; Wang et al., 2019f; Hu et al., 2018; Sha et al., 2019)
	KKBox		(Xin et al., 2019; Wang et al., 2019g)
Co-Curricular Interest	*Yelp*		Shi et al., 2015; Shi et al., 2018; Catherine & Cohen, 2016; Yu et al., 2013a; Yu et al., 2014; Wang et al., 2020b, 2019 g; Sha et al., 2019; Hu et al., 2018; Sun et al., 2018b; Zhao et al., 2017
	CEM		(Dadoun et al., 2019)
	Dianping-Food		(Wang et al., 2019c)
	DBLP, Meetup		(Luo et al., 2014)
	Sina-Weibo		(Wang et al., 2018a)

[14] Structured and Unstructured Data.
[15] https://www.techwyse.com/blog/searchengine-optimization/seo-efforts-to-getlisted-in-google-knowledge-graph/.
[16] schema.org.

assessment on different data benchmarks and remarkably sufficient deployment under different application scenarios proves their efficiency in recommendation (Guo et al., 2020). In this section, we categorize and discuss the utilized datasets and the accomplished works with respect to different application scenarios as summarized in Table 2. We categorize the benchmark-data in four categories (i.e., Readings, Shopping, Entertainment, and Matters of Particular Interests) and discuss their utilization from the perspectives of citing works. We also present an overview of the approaches that constructed their local subgraphs for experiments, either only from the input datasets or through the incorporation of external KGs. Moreover, we provide a descriptive summary of the KGE techniques and information mapping at the end of this section.

3.1. Categorization

In this section, we formally categorize the benchmark datasets in the following four categories.

3.1.1. Readings

In the reading's recommendation, Books, Articles, Blogs, News, etc., recommendations are well-known and imperative application domains. In case of Books where Book-Crossing (Ziegler et al., 2005), Amazon-Book (McAuley et al., 2015) and DB-book2014 (Di Noia et al., 2014; Cao et al., 2019; Song et al., 2019a) are data benchmarks, whereas Douban-Book (Yin et al., 2012), Intent-Books (Uyar & Aliyu, 2015) and LibraryThing[17] (Smith, 2007; Kanyundo & Kapondera, 2016) are also commonly utilized datasets. Informally, Douban-Book is crawled from Douban (ZHOU & ZHANG, 2013) that contains attributes like author's and publisher's info, publication year and user-item interaction data. Moreover, (Shi et al., 2018) utilized Douban-Book to construct their user-item subgraph, without any external KG. Similarly, rest of the four mentioned datasets contain users-books connectivity in binary feedback form. The subgraph of each dataset is created by mapping the corresponding books to their conforming entities in external KGs like DBpedia (Piao & Breslin, 2018; Cao et al., 2019; Song et al., 2019a), Freebase (Huang et al., 2018; Wang et al., 2019f; Qu et al., 2019) and Satori (Tang et al., 2019; Wang et al., 2019b; Wang et al., 2019c; Wang et al., 2019d; Wang et al., 2019e; Qu et al., 2019; Li et al., 2019c; Zhang et al., 2016a). While, LibraryThing contains the information of books, users and user ratings in range from 1 to 10; (Palumbo et al., 2020) utilized LibraryThing and Dbpedia to construct their knowledge graph.

On the other hand, News recommendation gave a challenging time to the researchers and still the challenges persist. According to (Wang et al., 2018b), it is challenging due to a few reasoning facts like requirement of common sense to understand, condensed nature of contents, lack of integrity, source authenticity and time-span sensitivity, etc. Therefore, (Wang et al., 2019b; Wang et al., 2019d) introduced KG to the scenario to improve the performance of recommendations through logical relations from people to news and news to news. Formally, Bing-News, i.e., MIND[18] - a collection of server-side logs of Microsoft news[19] - contains news titles and user interaction information like clicks, visits, etc., is considered as a benchmark dataset in this scenario. Microsoft Satori is commonly used to create subgraphs from this dataset (Qin & Liu, 2013). Moreover, Joseph & Jiang (2019) introduced a human annotated dataset and termed that as CNRec[20].

3.1.2. Online Shopping

With respect to the e-commerce, it is the most significant application scenario where Amazon-products (McAuley et al., 2015) is the benchmark dataset, and Alibaba's Taobao (Zhu et al., 2018; Zhu et al., 2019) is also a well-known dataset concerning the product's affair recommendation. Amazon-products contains user information like interaction records, behaviors, reviews, preferences, ratings etc. as well as item information such as category, type, description, etc. With the help of Freebase, (Ma et al., 2019) enhanced item's information and created item's graph from Amazon-products. Without the support of any external knowledge graph, (Ai et al., 2018; Xian et al., 2019; Zhao et al., 2017; Zhao et al., 2019b; Zhang et al., 2018c) utilized Amazon-products to construct their user-item subgraphs. Similarly, Alibaba's Taobao is also used by (Ye et al., 2019; Zhao et al., 2019b) for their experiments.

3.1.3. Entertainment

Currently, movie recommendation is among the most frequently triggered application scenarios. Movie-Lens (Harper & Konstan, 2015) and Douban-Movie (Wang et al., 2016) are popular datasets in this scenario. In Movie-Lens – the benchmark dataset, there are many stable datasets, e.g., MovieLens-100 K, 1 M, 10 M, 20 M, 25 M, etc., are collected from MovieLens[21]. The ml-latest-small is the latest version of this series. These datasets are comprised of ratings, tags and movies attributes. Whereas, Douban-Movie, crawled from Douban[22], is comprised of user information, social relations and movies attributes. Formally, the subgraphs are constructed in many ways in this scenario. For instance, (Piao & Breslin, 2018; Wang et al., 2019b; Wang et al., 2019c; Wang et al., 2019d; Wang et al., 2019e; Tang et al., 2019; Qu et al., 2019; Li et al., 2019c; Huang et al., 2018; Zhang et al., 2016a; Cao et al., 2019; Xin et al., 2019; Palumbo et al., 2018; Yang et al., 2018) constructed item graphs with the help of DBpedia, CN-DBpedia, Satori, Freebase, and IMDB[23], etc., by extracting relevant information from the mentioned KReps to enrich item's representation. Similarly, (Catherine & Cohen, 2016; Sun et al., 2018b; Wang et al., 2019g; Huang et al., 2019b; Song et al., 2019a; Sha et al., 2019; Yu et al., 2013a; Yu et al., 2013b; Yu et al., 2014; Palumbo et al., 2017) also used external KGs with Movie-Lens to construct subgraphs. Moreover, considering user ratings as a user-item relation and introducing it to the KG is also an approach of subgraph construction. Therefore, (Hu et al., 2018; Shi et al., 2015; Shi et al., 2018) constructed user-item subgraphs by directly exploiting users' interaction-data and movies' attributes in the datasets of Movie-Lens and Douban-Movie.

Like movies, music also belongs to the entertainment – a popular application scenario of recommendation. Last-FM (Schedl, 2016) and KKBox (Huang et al., 2019c; Chen et al., 2018) are popular datasets of this scenario. Last-FM – the benchmark dataset – is collected from last.fm[24] contains users and their previous music listening chronicles; whereas KKBox, provided by WSDM cup 2018 challenge[25], comprises music description and user interaction data. Formally, (Wang et al., 2019c; Wang et al., 2019d; Wang et al., 2019e; Wang et al., 2019f; Huang et al., 2018) and (Song et al., 2019a; Sha et al., 2019) constructed item graphs and user-item graphs respectively, by extracting relevant information from Satori and Freebase. However, (Hu et al., 2018) directly constructed user-item graph from Last-FM and, (Xin et al., 2019) and (Wang et al., 2019g) constructed item graph and user-item graph respectively, from KKBox without the support of any external knowledge repository.

3.1.4. Matters of particular interest

It is quoted *"Recommendation about interests, attracts!"* because recommendations of the choices of particular interests attract people. It is why Matters of Particular Interest (MoPI) is an interesting application

[17] https://www.librarything.com/.
[18] MIcrosoft News Dataset, https://msnews.github.io/.
[19] https://microsoftnews.msn.com/.
[20] Content based News Recommendation.
[21] https://movielens.org/.
[22] Network of Chinese social media, https://movie.douban.com/.
[23] https://www.imdb.com/.
[24] https://www.last.fm/ & https://millionsongdataset.com/lastfm/.
[25] https://wsdm-cup-2018.kkbox.events/.

scenario of the recommendation processes. Specifically, Yelp (Zhang & Pan, 2014; Asghar, 2016) – the data benchmark, CEM[26] (Dadoun et al., 2019) and Dianping-Food (Zhang et al., 2013; Li et al., 2019a) are concerned datasets in this scenario. Yelp is comprised of users, their check-ins, reviews and business tags; whereas CEM and Dianping-Food contain information about trips and restaurants respectively. Moreover, (Hu et al., 2018; Zhao et al., 2017; Shi et al., 2015; Shi et al., 2018; Catherine & Cohen, 2016; Sun et al., 2018b; Sha et al., 2019; Yu et al., 2013a; Yu et al., 2014) and (Wang et al., 2019f) constructed user-item graphs and item graph respectively, from Yelp, (Dadoun et al., 2019) used CEM for next trip recommendation, and (Wang et al., 2019c) utilized Dianping-Food for restaurant recommendations.

Online applications send different recommendations to the users on social media is a social application scenario of recommendation for users. For instance, DBLP[27], Meetup[28] and Sina-Weibo[29] are well-known datasets in this scenario. Formally, DBLP is used to send recommendations about academia, research, and recent issues of conferences and journals to the researchers. Meetup recommends offline (i.e., face to face) meetings with the people of potential interest, and Sina-Weibo suggests about the adored celebrities to the users through their previous interaction-data on social media. Moreover, (Wang et al., 2018a) exploited Satori with Sina-Weibo to construct their local item graph.

3.1.5. Discussion

Although in major concern we discussed such approaches that exploited single and multiple datasets in their experiments, there are few more latest studies that incorporated different varieties of application scenarios with respect to the used datasets into their experimental work. In KG-based recommendation generation, many current researchers utilized more than a single datasets to evaluate the performance of their proposed approaches. For example, (Yang et al., 2020) applied Last-FM, MovieLens-1 M and Book-Crossing, (Palumbo et al., 2020) exploited Last-FM, MovieLens-1 M and LibraryThing, and (Wang et al., 2020b) utilized Last-FM, Amazon-Book and Yelp, etc. Moreover, machine-based Question-Answering (Park et al., 2021; Zhang et al., 2020a; Qin et al., 2019; Huang et al., 2019b; Zheng & Zhang, 2019) is another emerging application scenario as well as a hot research area in KG-based recommendation methods now a days.

3.2. KG Embedding and information mapping

KGE is an efficient semantic illustration technique for representing graphs (i.e., entities and relations) in the continuous embedding (i.e., low dimensional vector) space. Various translators are introduced to convert KGs to the embedding space. For instance, TransE (Bordes et al., 2013) translates KG triplets (i.e., e_h, r, e_t) to the vector space as $e_h + r \cong e_t$, where e_h and e_t show the head and tail entities respectively and r is the relation between them. (Wang et al., 2014) mapped the embedding vectors of e_h and e_t to the r-based hyperplane. It aimed to project any e_h or e_t to various projection vectors with respect to different relations. TransR (Lin et al., 2015) extended the concept of r-based hyperplane to r-based influential space. TransA (Xiao et al., 2015a) is a metric learning-based KGE approach, and TransG (Xiao et al., 2015b) is a generative model-based KGE technique proposed to alleviate the problem of multiple-semantics (i.e., more than one meaning) of a given relation. (Grover & Leskovec, 2016) introduced node2vec algorithm to learn continuous representations of features of entities in the graph. They mapped the entity representations to the low-dimensional features' space to maximize the likelihood of stabilizing the graph hierarchy of neighboring nodes. TransA (Jia et al., 2016) is a locally adaptive KGE technique used to determine optimized loss function through the adaptive estimation of its boundary-span over different KGs. ETransR (Lin et al., 2017) automatically modeled the features' representations of entities and relations in the continuous embedding space to quantify the semantic relevance of information instances to be used for knowledge resolution. (Wang et al., 2017) enriched KG embeddings through a Patient Disease Medicine (PDM)-KG and used the enhanced representations for medicine recommendation to the patients. PrTransH (Li et al., 2019b) used an improved probabilistic approach to learn KG-embedding representations for link prediction in the medical domain.

(Catherine & Cohen, 2016) mapped graph entities (i.e., words or phrases) to KBs and claimed that their mappings are straightforward since the KBs involved in their paper are based on structured data. However, wikifiers or entity-linkers are essential in case of generic KBs (e.g., YAGO, Wikipedia, NELL, etc.). (Yang et al., 2016) introduced to map fresh entities and fresh relation-types into the current embedding representations without the necessity of retraining the model. (Bellini et al., 2018) accessed freely available entity mappings from DBpedia knowledge repository to enhance the information sets of Movie-Lens 20 M, Amazon Digital Music and LibraryThing datasets. (Li et al., 2020) introduced a quadruplet (s, v, o, p) – where s, v, o and p denote subject, verb (predicate), object and property respectively – to characterize the medical data in the KG, used PrTransH to translate their KG to the embedding space, and feasibly mapped the embedding representations to ICD-10[30] (Brämer, 1988) to enhance the knowledge base. (Chen et al., 2020a) enhanced information instances of the Amazon-Book dataset through reliably mapping their embedding representations to the conforming entities (via title matching) in Freebase – the external knowledge repository – for their experiments. (Khan et al., 2022b) enriched entity information of the underlying datasets (i.e., Amazon-Book, Last-FM and Bing-News) by effectively mapping their representations to the relevant entities in the external knowledge repositories (MS-Satori, DBpedia-ontology[31], MS-Satori respectively).

4. Methodology

In this section, we discuss the implementation techniques of KG-based recommendation approaches with respect to different application scenarios. Formally, we categorize this Section (i.e., Section 4) as follows: In Section 4.1 (i.e., Preliminaries), we introduce the basic concepts and notations of KG-based information processing. In section 4.2 (i.e., KG Exploitation at Information Level), we describe the utilization of KG-based side information in real world data processing and management systems. In section 4.3 (i.e., KG-based Recommendation Methods), we categorically discuss KG-based recommendation methods in details. In Section 4.4 (i.e., Fairness Aware Recommendation Methods), we explain fair and biased recommendations, and discuss fairness aware recommendation approaches. And in Section 4.5 (i.e., Discussion), we present the summary of Section 4.

4.1. Preliminaries

We introduce the fundamental concepts and notations to be further utilized in the context of this paper, basics of the KG-based recommendation methods, types of local subgraphs, and information refinement policies; in this section.

4.1.1. Concepts and notations

A directed graph G *such that* $G = (V, E, \Phi, \Psi)$, where V is the set of entities,E is the set of relations between entities in V as $E \subseteq \mathrm{V} \times \mathrm{V}$, Φ is a type mapping function f for entities as $\Phi_f : V \rightarrow A$, and Ψ is a type mapping function f for edges as$\Psi_f : E \rightarrow R$. Each entity $e \subseteq V$ is mapped to

[26] Customer Experience Management – an Amadeus database.
[27] Digital Bibliography and Library Project, https://dblp.uni-trier.de/xml/.
[28] https://www.meetup.com/.
[29] https://weibo.com/.
[30] International Classification of Diseases, 10th Revision.
[31] https://wiki.dbpedia.org/services-resources/ontology.

a particular entity type in A such that $\Phi_f(e) \in A$, and each edge $r \in E$ belongs to a particular relation type in R such that $\Psi_f(r) \in R$. Thereby, if $|A|\rangle 1$ or $|R|\rangle 1$; the G is heterogeneous otherwise it is homogeneous (Fu et al., 2017). Moreover, a directed subgraph $\mathscr{G}$ such that $\mathscr{G} = \{(V, E) \mid \mathscr{G} \in G\}$ where V shows the set of nodes or entities and E represents the set of edges or paths or relations between entities. Typically, entities interlinked via relations create triplets (i.e.,$< e_h, r, e_t >$ where e_h and e_t represent the head and tail entities respectively, and r denotes the relation or path $\mathscr{P}$ between them), and $\mathscr{G}$ exploits the triplets to store and process information.

Knowledge Graph Embedding (KGE) is the process of embedding (converting, translating) $\mathscr{G}$ to its low-dimensional vector space (n-dimensional vector form) to preserve it's information in the machine interpretable form. KGE provides effective solutions to the challenges of graph-analytics[32] (Cai et al., 2018). The process of KG construction and graph/entity embedding is represented in Fig. 2.

Meta path is a collection of paths $\wp$ denoted by $\wp$. A $\wp$ from e_a to e_d is $\wp(e_a, e_d)$; identical to $\wp = \{r_1, r_2, r_3\}$ as $\wp = e_a \xrightarrow{r_1} e_b, e_b \xrightarrow{r_2} e_c, e_c \xrightarrow{r_3} e_d$. On joining the triplets, it implies that $\wp \Rightarrow e_a \xrightarrow{r_1} e_b \xrightarrow{r_2} e_c \xrightarrow{r_3} e_d$, where $\{r_1, r_2, r_3\}$ are the consecutive relations. Similarly, Meta graph is a collection of Meta paths. It can represent the portions of $\mathscr{G}$ as a part or whole in its depiction (Fang et al., 2016).

Hop-neighboring length is the length of Meta Path $\wp$ between the target entity e_t and the head entity e_h denoted by $\mathscr{H}^L_{e_h}$, where L represents the number of relations in $\wp$ or the path-distance between e_h and e_t. In $\wp = e_a \xrightarrow{r_1} e_b \xrightarrow{r_2} \cdots \xrightarrow{r_k} e_k$, e_k is k-hop neighbor of e_a; and $e_a \in \mathscr{H}^0_{e_a}$, $e_k \in \mathscr{H}^k_{e_a}$ and $e_t \in \mathscr{H}^L_{e_a}$ show *0-hop*, *k-hop* and *L-hop* neighboring lengths of e_h with e_a, e_k and e_t respectively.

The sets of m users and n items are expressed as $\mathscr{U} = \{u_1, u_2, \cdots u_m\}$ and $\mathscr{V} = \{v_1, v_2, \cdots v_n\}$ respectively, and user-item interaction matrix $\mathscr{M} \in \mathbb{R}^{m\times n}$ is defined as $\mathscr{M}_{i,j} = [1 \quad 0]$, where 1 denotes the occurrence of user-item interaction and 0 shows no evidence. Moreover, user-item relevance is obtained from $\mathscr{M}$ based on their previous interaction record. It is expressed as $\mathscr{R}^k_u = \{e_t | (e_h, r, e_t) \in \mathscr{G} \text{and} e_h \in \mathscr{R}^{k-1}_u\}$, where $\mathscr{R}^k_u$ is user-item relevance shown through the set of k-hop neighboring sequence. A user u_i prefers item v_j or not is a probabilistic situation defined via the following equation.

$$\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j) \tag{1}$$

Where $\check{\mathscr{Y}}_{(i,j)}$ shows the user preferences and $f()$ is used to convert latent vectors of user-item embedding to scalar-valued list (i.e., $\check{\mathscr{Y}}_{(i,j)}$ is sorted in descending order).

Explicit path embedding is used to interlink pairs of u_i & v_j to directly model the relations between them in $\mathscr{G}$. Let, there are $\mathscr{K}$ paths in $\mathscr{G}$ to interlink u_i and v_j, their interaction is defined via $\mathscr{E} = g(\mathscr{E}_p)$, where $g(\,)$ summarizes the information acquired from each $p|p = 1, 2, \cdots, \mathscr{K}$ of path embedding and $\mathscr{E}_p$ is path p's embedding. Therefore, the user preference is calculated via $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathscr{E}_I)$, where $\mathscr{E}_I$ shows the Ultimate Interaction Embedding (UIE).

Frobenius Norm or Euclidian Norm is a matrix norm of matrix A of the form $m \times n$, defined as "square root of sum of absolute squares of its elements" and mathematically expressed as:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2} \tag{2}$$

Where F is Frobenius Norm, and i and j are elements of A (Van Loan, 2000; Moler & Van Loan, 2003). F is equivalent to the square root of matrix trace AA^{T_C}, where A^{T_C} is conjugate transpose of A; thus, written as $\|A\|_F = \sqrt{\mathrm{Tr}(AA^{T_C})}$. F is also used for vector D^2-norm, and known as vector norm. We summarized the preliminary concepts and notations in Table 3.

4.1.2. Basics of KG-based knowledge utilization

Many exertions are made to provide general semantic descriptions and representations essential to formally define the characteristics of KG or Linked Open Data (LOD); however, there exist no such formal definition that is widely accepted (Ji et al., 2021). KG obtains the required information from different information sources (KReps), integrates it in the ontologies and applies reasoning mechanisms to infer new knowledge based on the existing facts and figures (Ehrlinger & Wöß, 2016). Intuitively, numerous approaches have incorporated the concept of utilization of hierarchical information into the KGs. For instance, (Leacock & Chodorow, 1998; Ponzetto & Strube, 2007) used KG-derived hierarchical information to build inter-concepts relatedness, (Morin & Bengio, 2005) introduced hierarchical information-based neural-language modelling framework into the WordNet. (Hu et al., 2015) learned entity-representations from entity hierarchies of Wikipedia and stimulated a number of researchers (Krompaß et al., 2015; Xie et al., 2016) to incorporate the required hierarchy structures from KReps into the KG to enhance the underlying representations. Entity-based hierarchical information has proven that it is effective in model's enhancement strategies through the enhancement of the underlying information being utilized by the model (Hu et al., 2015; Xie et al., 2016).

Current approaches exploit embedding based information incorporation. To acquire zero-shot fine-grained-based named-entity typing from the underlying information, (Ma et al., 2016) utilized hierarchical information for prototype-driven label embedding. (Cao et al., 2020) claimed that the current methods only provide the missing values to the entities or complete the missing/unreadable relations among the entities during the enrichment of KGs. Also, the current approaches provide no attention to the long-tail entities in the KGs. Therefore, they proposed an end-to-end complete approach for KG enrichment that forecasts the missing attributes (i.e., entity, relation, property) and extracts the relevant information from the open-web to provide it to the long-tail entities in the KG.

In the literature, different techniques are available that are introduced and exploited to enrich the semantic representations of the underlying data in KGs. For instance, (Ji et al., 2017) introduced the enrichment of entity description via sentence-level Attention Mechanism (AM) with Piecewise Convolutional Neural Network (PCNN). HATT (Han et al., 2018) applied structure-based selective AM by appending the attentive representations of each next hierarchical layer to obtain the relational hierarchy of information in the target data source. Moreover, (Zhou et al., 2016) used word-level AM with BI-LSTM technique and (Soares et al., 2019) utilized deep-transformers to learn the pre-trained relation-based representations. TKMF (Jiang et al., 2019) uses relevant words, topics and sentences to enrich the sentence-level representations. Furthermore, (Monnin et al., 2019) proposed entity-based knowledge reconciliation framework to examine the mutual similarity, generality and relevancy among the nodes across the KG. From the perspectives of language models' processing, this work can be considered as the extension of entity resolution in object matching. In this study, they investigate how GCN[33] can be utilized, how its performance can be assessed in experiments on real world data, and how the variations of genes influence the drug responses. (Alam et al., 2017) suggested that reconciled KGs can be used to summarize text and to detect specific knowledge in it. They proposed KG reconciliation approach based on two differently described documents of text about few identically happened events. The event-reconciliation is

[32] How to find, explore and utilize the information hidden in knowledge graphs.

[33] Graph Convolutional Networks.

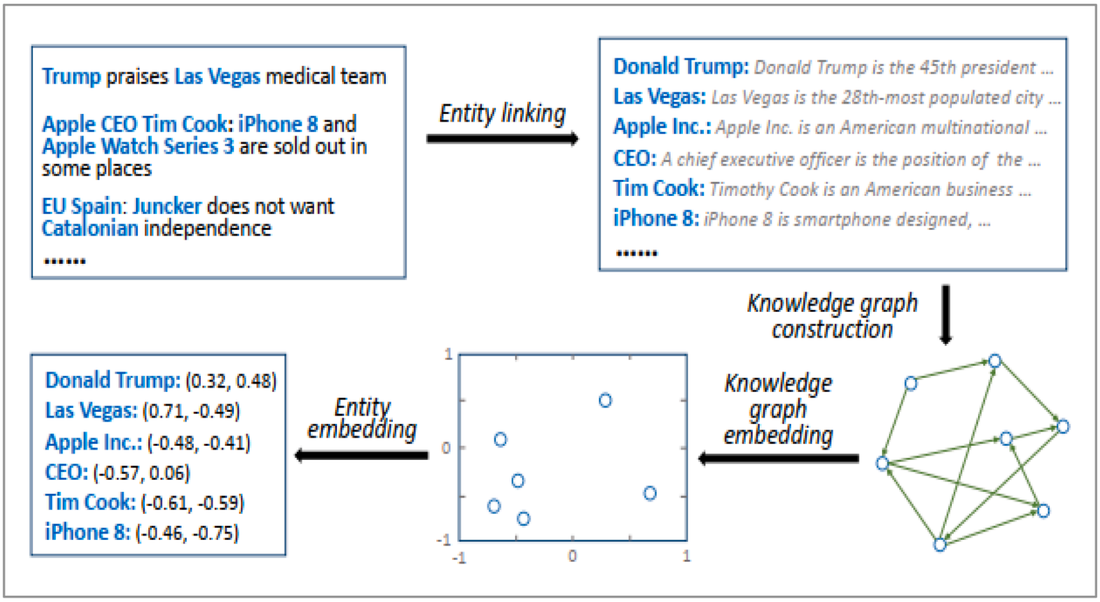

Fig. 2. Entity identification in the text, entity inter-linkage, knowledge graph construction, knowledge graph embedding, and entity embedding based on the purified/modelled text data with respect to a specific application scenario (Wang et al., 2018b).

Table 3
Notations and their brief description. Abbreviations used – D: Dimensional, LV: Latent Vector, LS: Latent Space, MP: Meta Path, HN: Hop Neighboring, SF: Scoring Function, MF: Metric Function, RS: Results Store.

Notations	Brief Explanation
$\mathscr{u}_i \& \mathscr{v}_j$	User i & Item j
$\mathscr{U} = \{\mathscr{u}_1, \mathscr{u}_2, \cdots \mathscr{u}_m\}$	$\mathscr{u}_i$'s set of m users
$\mathscr{V} = \{\mathscr{v}_1, \mathscr{v}_2, \cdots \mathscr{v}_n\}$	$\mathscr{v}_j$'s set of n items
$\mathscr{p} \& \mathscr{p}_l$	Path & Path covering one relation l like r_l
$\wp(\mathscr{e}_w, \mathscr{e}_z) = \{r_1, r_2, \cdots, r_q\}$	MP: $\wp = \{r_1, r_2, r_3\}$ as $e_w \xrightarrow{r_1} e_x$, $e_x \xrightarrow{r_2} e_y$, $e_y \xrightarrow{r_3} e_z$
Theta Θ	Set of weights of Meta Paths $\wp$ s
$\mathscr{H}$: $\mathscr{H}^l_{e_a}$	HN: l is distance from e_a to tail entity.
$\mathsf{u}_i \in \mathbb{R}^{1\times d}$	$\mathscr{u}_i$'s LV with 1-D LS
$\mathbf{U} \in \mathbb{R}^{m\times d}$	$\mathscr{U}$'s LV with m-D LS
$\mathsf{v}_j \in \mathbb{R}^{1\times d}$	$\mathscr{v}_j$'s LV with 1-D LS
$\mathbf{V} \in \mathbb{R}^{n\times d}$	$\mathscr{V}$'s LV with n-D LS
$r_l \& \mathbf{r}_l \in \mathbb{R}^{1\times d}$	A relation $r_l = e_a \xrightarrow{l} e_b$ & its LV with 1-D LS
$\mathscr{e}_k \& \mathbf{e}_k \in \mathbb{R}^{1\times d}$	An entity k & its LV with 1-D LS
$\check{\mathscr{Y}}_{(i,j)}$	SF: Predicts $\mathscr{u}_i$'s preference towards $\mathscr{v}_j$
$\partial(\,) \& \mathscr{D}$	MF & RS of distances calculated by $\partial()$
$x : Y \rightarrow Z$	x maps each object from Y to any object in Z
$\mathscr{E}_{\mathscr{p}} \& \mathscr{E}_I$	Path embedding & ultimate interaction embedding

accomplished via KG-structure-based similarities using RDF2Vec (i.e., KG embedding) and semantic roles-based subsumption of knowledge hierarchies.

Although learning entities and relations from embedding space acquired valuable popularity and performance on benchmark datasets with respect to some applications scenarios, they fail to effectively model higher-order paths in the complex relations. Reasoning over higher-order relations needs the information of specific paths among the connecting entities. Random walk sampling have been widely exploited for this purpose by path-based methods. For instance, the famous Path Ranking Algorithm (Lao & Cohen, 2010) selects a relational-path based on the combination of applied path-constraints and perform max-likelihood classification to rank it. (Gardner et al., 2014) improved path-search technique via introducing vector-space relevancy heuristics into the random walk strategy by integrating textual contents into the vector space. Moreover, (Neelakantan et al., 2015) recursively applied compositionality technique to constitute inference-based relational paths via RNN model. (Das et al., 2016) proposed chain-of-reasoning to empower numerous reasons and demonstrated logical compositions among entities and relations via AM in the text. Recently, (Chen et al., 2018) proposed an aggregated Bi-inferencing Variational Framework (BVF) to accomplish the tasks of higher-order path-finding and path-reasoning concurrently.

KR-EAR (Lin et al., 2016) proposed novel knowledge representation (KR) model based on KG entities, attributes and relations (EAR); and further divided current KG-relations in two subtypes, i.e., relations and attributes. The triplet-based relationship has two types (i.e., relation-based triplets and attribute-based triplets) with respect to the KG (Sun et al., 2017). Typically, relation-based triplets denote the relation between entities (e.g., Barack, is-the-first-name-of, Obama), and attribute-based triplets literally represent an attributed-value of the concerned entity (e.g., Barack, born-in, 1961). KG-based data (i.e., entities, relations, attributes and descriptions) is greatly applicable in various application scenarios regarding NLP[34] tasks (Sun et al., 2018a). About large Kreps (e.g., DBpedia, Wikidata, etc.), the concept "*it is nearly impossible to identify which attributes are important and which are not*"; arisen the terminology of attribute selection. (Sun et al., 2018a) exploited textual information from the external users (e.g., comments, reviews, randomly generated text-data, etc.) to assess the relative significance of attributes of a specific entity. Actually, they used *token* and *sub-token*-based embedding to compare attributes (i.e., name, value, etc.) of the concerned entities with the external text data; and ranked the underlying attributes accordingly. (Gutlein et al., 2009) proposed LFAS[35] strategy to reduce the numbering expansion of the available attributes in each next iteration of the forward selection phase. (Wu & Wang, 2018) proposed a novel KGE approach to jointly model the relations among entities and the numeric attributes of the relations.

4.1.3. Basics of KG-based recommendation

In this section, we categorically discuss the research-work proposed by different researchers in KG-based recommendation. A distinguishing benefit of KG-based recommendation over traditional RSs is it provides self-explanatory reasoning with suggestions to the end users.

[34] Natural Language Processing.

[35] Linear Forward Attribute Selection.

Explainable recommendation is an attracting plus-point of KG-based recommendation systems. When RSs provide clear explanations with their suggestions, users feel satisfaction in selection of their required items (Zhang & Chen, 2018).

We consider KG-based recommendation in path-based methods, embedding-based methods and unification-based methods (i.e., a possible combination of path and embedding-based methods). Path based methods create/use user-item interactions' KG and utilize similarity-based channel of entity-connection for recommendations. Path based methods crawl (i.e., perform random-walk) over the graph and enumerate relations among paired entities with respect to the given condition to construct a feature-matrix. A binary classification technique[36] is learned on previously created feature-matrix to infer the missing entities and connections in the KG (Jagvaral et al., 2020). In short, these methods exploit the sequence of previous user-item interactions to generate new recommendations based on the similarities of paths between the sequences of previous and new interactions. To calculate path-based similarities among entities (i.e., users and users, users and items, and items and items) in the graph, various similarity measures/coefficients proposed by different researchers are available in the literature. Therefore, we one-by-one introduce some of the very well-known and commonly used similarity calculation methods in the following segment.

4.1.3.1. Jaccard similarity Coefficient. Jaccard Similarity Coefficient (JSC) primarily basis on explicit ratings of users about the items. JSC is the ratio of the cardinality of mutually identical ratings to all ratings with respect to the concerned users. Mathematically, it is defined as:

$$JSC(\alpha,\beta) = \frac{(\mathscr{V}_\alpha \cap \mathscr{V}_\beta)}{(\mathscr{V}_\alpha \cup \mathscr{V}_\beta)} \tag{3}$$

Where $\mathscr{V}_\alpha$ and $\mathscr{V}_\beta$ show the rated cardinalities of the items, and α and β denote two different users.

4.1.3.2. Tanimoto similarity Index. Tanimoto Similarity Index (TSI) – also known as Jaccard Similarity Index (Dehmer & Varmuza, 2015) – is a measure of graph-data similarity. It is commonly used to find the similarity between the connected subgraphs. For two isomorphic and interlinked subgraphs α and β, TSI is defined as:

$$TSI(\alpha,\beta) = \frac{\sum_{s=1}^{|Z|} \cap \left(\gamma_s^\alpha, \gamma_s^\beta\right)}{\sum_{s=1}^{|Z|} \cup \left(\gamma_s^\alpha, \gamma_s^\beta\right)} \tag{4}$$

Where γ shows the information instances and Z is the set of embedded *meta*-graphs of α and β.

4.1.3.3. Adamic-Adar similarity Index. Adamic-Adar Similarity Index (AASI) takes direct neighbors as well as neighbors of the neighbors into account to calculate similarity between users and users or items in the graph. It assumes that two users are similar if they possess more neighbors in common and their common neighbors possess fewer direct neighbors. Mathematically, AASI between the users α and β is defined as:

$$AASI(\alpha,\beta) = \sum_{z\in(\Gamma(\alpha)\cap\Gamma(\beta))} \frac{1}{\log|\Gamma(z)|} \tag{5}$$

Where Γ shows the set of neighbors.

4.1.3.4. PathSim similarity technique. PathSim is the measure of Meta-path-based inter-entity similarity (Sun et al., 2011). It is commonly used by path-based recommendation methods. PathSim between two identical entities (i.e., user and user, user and item, and item and item) α and β is calculated is:

$$PathSim(\alpha,\beta) = \frac{2\times|\{\mathscr{P}_{\alpha\to\beta} : \mathscr{P}_{\alpha\to\beta} \in \wp\}|}{|\{\mathscr{P}_{\alpha\to\alpha} : \mathscr{P}_{\alpha\to\alpha} \in \wp\}| + |\{\mathscr{P}_{\beta\to\beta} : \mathscr{P}_{\beta\to\beta} \in \wp\}|} \tag{6}$$

Where $\mathscr{P}_{\alpha\to\beta}$, $\mathscr{P}_{\alpha\to\alpha}$ and $\mathscr{P}_{\beta\to\beta}$ show the exemplary path instances between the entities α and β, α and α, and β and β respectively.

4.1.3.5. Cosine similarity measure. Cosine Similarity (CosSim) is measure of the angle between two explicitly rated vectors. The smaller is the angle between the vectors; the greater is the similarity between them and vice versa. The measure of angle ranges from 0 to 1 and the value of angle indicates the extent of similarity between user α and user β. Mathematically, the CosSim is defined as:

$$CosSim(\alpha,\beta) = \frac{\sum_{\mathscr{v}\in\mathscr{V}(\alpha,\beta)} r(\alpha,\mathscr{v})\cdot r(\beta,\mathscr{v})}{\sqrt{\sum_{\mathscr{v}\in\mathscr{V}(\alpha,\beta)} r(\alpha,\mathscr{v})^2}\cdot\sqrt{\sum_{\mathscr{v}\in\mathscr{V}(\alpha,\beta)} r(\beta,\mathscr{v})^2}} \tag{7}$$

Where $\mathscr{V}(\alpha,\beta)$ is the cardinality of the items mutually rated by the users α and β, $\mathscr{v}$ is an individual item, and $r(\alpha,\mathscr{v})$ is the measure of the explicit rating of $\mathscr{v}$ given by α.

Translation-distance and semantic-matching techniques[37] are commonly used to embed KG to its vector (low-dimensional) representations. For instance, TransD (Ji et al., 2015), TransE (Bordes et al., 2013) and TransR (Lin et al., 2015) are translation-distance; and DistMult (Yang et al., 2014) is semantic-matching technique. Embedding-based methods are known as regularization-based methods if regularization terms are imposed on loss function to learn entity-embedding in the KG structure (Zhang et al., 2016a; Cao et al., 2019).

4.1.4. Types of local subgraphs

With respect local subgraph construction and embedding, recommendation approaches are divided in two types, i.e., 'user-item-graph' and 'item-graph' embedding. In *user-item-graph* embedding, user-item graph depends on the paths among users and items in KG, and user-item information is embedding to the vector space based on the relations among them. The required references are calculated using user-item or user-item-relation information via $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j)$ or $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathbf{r})$ respectively. While in *item-graph* embedding, subgraph is constructed with the information of items and their associated attributes/contents. Item information is also extracted from the external KReps if required. Subgraphs are encoded to the Euclidian space to compactly preserve, comprehensively describe and efficiently utilized their information. The preference calculation methods can be varyingly applied with respect to the significance of relations' information.

In the consulted literature, three approaches relate to user-item-graph embedding, i.e., CFKG (Zhang et al., 2018c), DKFM (Dadoun et al., 2019) and SHINE (Wang et al., 2018a), and the rest of the papers belong to the item-graph embedding. Joint utilization of semantically interlinked data and semantic representation of users, items and the relations among them is possible via unification-based methods. These methods refine the representations of entities and/or relations based on the KG-relation structure, comply the concept of embedding propagation, and compute users' preferences via $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j)$.

4.1.5. Information refinement

Refining entity representations befalls in two categories, i.e. *refining user representations* and *refining item representations*. In *refining user representations*, the embeddings of past interacted items as well as the embeddings of past interacted item's multi-hop neighbors are utilized to learn the embedding of users. It is based on user's previous interaction with a confirmation that the user or its selected item is the head entity in

[36] decision tree, support vector machine or logistic regression.

[37] Both of these techniques have many other types besides the few we mentioned.

the corresponding triplet of that ripple set. For instance, $\mathfrak{S}^1_{u_i} = (e_h, r, e_t) \in \mathfrak{S}^k_{u_i}$ where every $e_h \in \mathfrak{S}^1_{u_i}$ in each corresponding $\wp$ represents user or its selected item. This is called the propagation of user preferences in KG because the propagation starts from user or its selected items in this process. Thus, following objective function is used to learn u_i's representations:

$$u_i = g_u\left(\left\{\mathfrak{S}^k_{u_i}\right\}^L_{k=1}\right) \tag{8}$$

Where $g_u()$ is used to concatenate multi-hop entity-embeddings with the bias variable.

Whereas in *refining item representations*, the embeddings of item's multi-hop neighbors are combined to refine item representations. To obtain $\mathbf{v}_j$'s representations, the following objective function is defined:

$$\boldsymbol{v}_j = g_v\left(\left\{\mathfrak{S}^k_{v_j} \odot \mathscr{H}^k_{v_j}\right\}^L_{k=1}\right) = g_v\left(\left\{(\mathfrak{S} \odot \mathscr{H})^k_{v_j}\right\}^L_{k=1}\right) \tag{9}$$

Where $g_v()$ shows the concatenation of item's multi-hop neighbor's embeddings.

The multi-hop neighbors' concatenation is completed in two steps: First, entity's k-hop neighbors' representation is learned through $\boldsymbol{e}_{\mathfrak{S}^k_{v_j}} = \sum_{(e_h,r,e_t)\in\mathfrak{S}^k_{v_j}} \boldsymbol{\alpha}_{(e_h,r,e_t)}\boldsymbol{e}_h$, where $\boldsymbol{\alpha}_{(e_h,r,e_t)}$ describes different neighbors, $\boldsymbol{e}_h$ defines $\mathrm{Ag}(\boldsymbol{e}_h, \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}) | e_h \in \mathfrak{S}^k_{v_j}$, and Ag shows the aggregation operator. Second, the learned representations of k and *(k-1)*-hop neighbors of item are aggregated via *aggregators (*i.e.*, neighbor, sum, concat and bi-interaction aggregation operations)* to complete the concatenation process. At the completion of corresponding operation of each aggregator, a non-linear transformation is applied on the consequent process. Formally, *Neighbor* aggregator replaces entity representations with the neighbors' representation as $\mathrm{Ag}_{nei} = \Phi(W.\boldsymbol{e}_{\mathfrak{S}^k_{v_j}} + b)$, *Sum* adds the representations as $\mathrm{Ag}_{sum} = \Phi(W.(e_h + \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}) + b)$, *Concat* concatenates the representations via $Ag_{con} = \Phi\left(W.\left(e_h \oplus e_{\mathfrak{S}^k_{v_j}}\right) + b\right)$, and *Bi-interaction aggregator* calculates the product of element-element relations between entities, and then adds as $Ag_{bii} = \Phi\left(W.\left(e_h + \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}\right) + b\right) + \Phi\left(W.\left(e_h \odot \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}\right) + b\right) = \Phi\left(W.\left[\left(e_h + \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}\right) + \left(e_h \odot \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}\right)\right] + 2b\right)$, where $\Phi\left(W.\left(e_h \odot \boldsymbol{e}_{\mathfrak{S}^k_{v_j}}\right) + b\right)$ projects the information of similar entities more than those of dissimilar to the concatenation process. In this discussion, three approaches relate to *refining user representations*, i.e., RippleNet (Wang et al., 2019b), AKUPM (Tang et al., 2019) and RCoLM (Li et al., 2019c), and the rest belong to the *refinement of item representations*.

4.2. KG Exploitation at information level

*The utilization of KG data (*i.e.*, KG-based side information) for decision making through the real-world deployed systems is known as KG Exploitation at Information Level*. KG appeared as a promising technique to manage, provide, interlink, process, and employ information in the form of abundant entities and their mutual relationships in normal and embedding states for exploitation (Li et al., 2021a). KGs are greatly utilized in different real-life application fields (e.g., education, health, tourism, business, etc.,) due to their natural theme of information presentation and processing.

In this section, we discuss several latest studies that exploited KG information for different determinations in daily-based application scenarios. For instance, the webpage of La-Rioja tourism[38] is a well-deployed and popular example of KG construction and its information's exploitation for the promotion of tourism-industry in Spain (Alonso-Maturana et al., 2018). The information (e.g., accommodation, cuisines, wineries, routes, activities, etc.) is presented and maintained in RDF format in OWL. Existing ontologies (i.e., vocabularies like Route, OnTour, Harmonise, etc.) are extended and aggregated to construct the current underlying KG via the implementation of Digital Semantic Model (DSM). Currently, the KG has 7 K + contents comprised of 675,368 triplets, 472,361 relations and 67,284 entities. The information exploitation is made possible by linking the KG with search engines to explore and utilize its knowledge objects, information and visualization systems via unification of map and geo-positioning. Moreover, in European Union (EU), public procurement is a Facilitation Centre that provides different facilities to the concerned communities under the government's supervision. In this perception, based on the procurement KG, (Soylu et al., 2020) developed a public-welfare platform to provide general assistance to the people. The platform contains ontologies and modulation APIs to perform different operation (i.e., EU-wide broadcasting, curation, integration, analysis, visualization, and multi-lingual-based data incorporation) to enhance the corresponding KG. Also, based on the stated KG, they developed a tool for information interfacing and communication of the end users.

(Shen et al., 2020) exploited KG information for entity masking and acquiring distractors for the already mask entities. They proposed DSR[39] technique and jointly optimized it with the language model of the masked entities. They concurrently used DSR for question-answering and KG-completion tasks. In this approach, incorporated KG-based structured information into the language models of raw text-data to learn their proposed model. Correspondingly, Asynchronous annotation of datasets is challenging for the information handling systems. Therefore, (Cardoso et al., 2020) proposed a Historical KG (HKG) with the following contributions: (1) Stored data of all previous versions of an ontology in a single KG, (2) Hence, decreased the data storage space, and (3) No data refinement is required. They proposed a structural mechanism to exploit the side-information of HKG in this work. Similarly, (Michel et al., 2020) proposed a project named "COVID-on-the-web" to provide experimental data about COVID19 to the current research community in KGs. The project acquires current information about COVID19 and enrich the dataset accordingly. The dataset of the project "COVID-on-the-web" (i.e., CORD19), contains 50 K + research articles with their complete text about the corona-virus. In major description, CORD19 has two KGs (i.e., Named Entity KG and Argumentative KG) that contain; (a) connected the available entities to the vocabularies of different external Kreps (like DBpedia, Wikidata, BioPortal, etc.), and (b) extracted the arguments to assist medical staff to analyze the trials necessary for decision making. To establish the required experimental environment, numerous evaluations and visualizations are performed to effectively exploit the concerned KGs.

Based on the side information of industrial KG (IKG), (Zheng et al., 2021) introduced Multi Agent Reinforcement Learning (MARL) approach to design and implement SCM[40] network. They constructed IKG and applied GNN-attributed embedding technique for semantic-based auto-configurable task determination and solution acquisition. (Wang et al., 2020c) modelled semantic-space via LSA[41] and semantic-path via Page-Rank algorithm to determine the potential semantic-paths in text-data with respect to the given sentences. The point worth mentioning is they considered a popular Knowledge Base (i.e., WordNet) as a data-benchmark in their study. (Wang et al., 2020a) proposed a

[38] https://lariojaturismo.com/en.
[39] Distractor Suppressed Ranking.
[40] Self-X Cognitive Manufacturing.
[41] Latent Semantic Analysis.

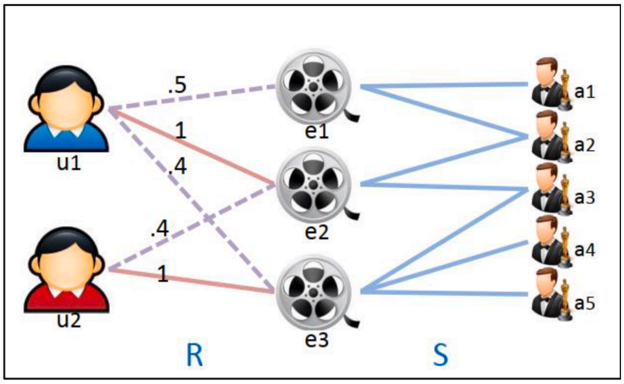

Fig. 3. The diffusion of users' preferences via item similarity matrix based on *meta*-path. S is Similarity Matrix and R is Feedback Matrix. Solid brown lines show the observed preferences and dashed connections signify the diffused preferences (Yu et al., 2013b).

novel substructure-based GRL[42] to subgraph-wise embed graph information to the Euclidian space. They deemed subgraphs as the substructures of KG; therefore, high correlations (mutual dependencies) must exist among the parts of KG. "*What is the extent of the mutual dependences*?" is the actual question. Thus, they developed Mutual Information (MI)-attributed substructure-based GRL approach and applied it on node and graph level information to maximize the ratio of MI. Earlier researchers (Basile et al., 2014) also agreed that KG-based data exploitation in recommendation can obviously beat the CF and content-based recommendation methods with respect to the performance. Similarly, LOD-empowered semantic-aware recommendation techniques have an upper edge in performance over the traditional RSs with ordinary (i.e., plan text-based) datasets (De Gemmis et al., 2015).

4.3. KG-based Recommendation methods

In this section, we informally categorize KG-based recommendation domain in *Matrix Factorization and Collaborative Filtering*, *Attention Mechanism*, *Deep Learning*, *Reinforcement Learning* based Methods and *KG-based General Recommendation Methods*; and discuss their implementation techniques in the respective categories. Also, we present tabularized comparative analysis of the techniques utilized in each category based on their applied *Methods*, *Venues*, *Frameworks*, *Datasets*, *Evaluation Parameters*, *Publication-year* and *References.*

4.3.1. Matrix Factorization (MF) and collaborative filtering (CF)-based methods

MF algorithms – a subclass of CF-based methods – decompose the underlying data (i.e., *user to item interaction matrix) to the product-operation of two low-dimensional quadrilateral matrices.* Although MF and CF are traditional and widely used recommendation methods, highly suffer from various technical limitations like cold start, data sparsity and gray sheep; while KG is efficacious to overcome these limitations. Therefore, various researchers exploited KG-based side information via MF and CF-based methods for recommendation. For instance, HeteRec (Yu et al., 2013b) enhanced user-item interaction-matrix $\mathscr{M}$ through the incorporation of Meta path-based similarities into the graph, and applied PathSim to calculate item-item similarity. Moreover, it declared different Meta paths, denoted by $\mathscr{D}$, to inter-link the users and items in the graph. It used $\mathscr{D}$ with PathSim to acquire $\mathscr{D}$ item-to-item identical matrices like $Sim^{(\wp)}$, where $Sim^{(\wp)} \in \mathbb{R}^{n\times n}$ and $\wp = 1, 2, \cdots, \mathscr{D}$. Further, it used $\widetilde{\mathscr{M}}^{(\wp)} = \mathscr{M}Sim^{(\wp)}$ to calculate $\widetilde{\mathscr{M}}^{(q)}$, where $\widetilde{\mathscr{M}}^{(\wp)}$ shows $\wp$-wise similarity-diffused and $\widetilde{\mathscr{M}}^{(q)}$ represents $\mathscr{D}$-diffused user preference matrices (i.e., $\overline{\mathscr{M}}_{u}^{\mathscr{D}(\wp)}$), as shown in Fig. 3. HeteRec applied non-negative MF (Ding et al., 2008) on $\overline{\mathscr{M}}_{u}^{\mathscr{D}(\wp)}$ to acquire different Meta path-based $\mathscr{D}$-refined user and item latent vectors. Mathematically, the process can be defined as:

$$\left(\check{\mathbf{U}}^{(\wp)}, \check{\mathbf{V}}^{(\wp)}\right) = \mathrm{argmin}_{\mathbf{U},\mathbf{V}} \|\widetilde{\mathscr{M}}^{(\wp)} - \mathbf{U}^{T}\mathbf{V}\|_{F}^{2}, \text{ where } \{\mathbf{UV}|\mathbf{U} \text{ and } \mathbf{V} \geq 0\} \tag{10}$$

Conclusively, preferences are unified with scoring function to obtain the recommendations as:

$$\check{\mathscr{Y}}_{(i,j)} = \sum_{\wp=1}^{\mathscr{D}} \theta_{\wp} \odot \check{\mathbf{u}}_{i}^{(\wp)T} \check{\mathbf{v}}_{j}^{(\wp)} \tag{11}$$

Where θ represents the weight of pair of user-item latent vectors and $\odot$ is the product operation.

Beside the general preference model, HeteRec-p (Yu et al., 2014) categorized users in $\mathscr{Z}$-categories based on their past behaviors and interactions; and generated customized recommendations. It aimed to calculate user-user similarity to recommend items. This approach believed that "*different Meta paths have different significance for different users*". Hence, the modified scoring function is defined as:

$$\check{\mathscr{Y}}_{(i,j)} = \sum_{\alpha=1}^{\mathscr{Z}} Sim_{cos}(\mathscr{C}_{\alpha}, \mathbf{u}_{i}) \sum_{\wp=1}^{\mathscr{D}} \theta_{\wp}^{\alpha} \odot \check{\mathbf{u}}_{i}^{(\wp)T} \check{\mathbf{v}}_{j}^{(\wp)} \tag{12}$$

Where $Sim_{cos}(\mathscr{C}_{\alpha}, \mathbf{u}_{i})$ is the cosine similarity between the target-user category $\mathscr{C}_{\alpha}$ and the concerned user $\mathbf{u}_{i}$, and $\theta_{\wp}^{\alpha}$ is the significance of Meta path $\wp$ with respect to the user category $\alpha|\alpha \in \mathscr{Z} \& \wp \in \mathscr{D}$.

Hete-MF (Yu eta al., 2013a) suggested to focus on $\mathscr{D}$-meta paths in KG to calculate item-item similarity in each $\wp$where$\wp \in \mathscr{D}$. It combined item to item regularization with weighted-non-negative MF method (Zhang et al., 2006) and enhanced latent vector representations of user to item interactions to improve the recommendation. Hete-CF (Luo et al., 2014) outperformed Hete-MF by calculating the potential interactions of users towards un-rated (i.e., unseen) items via the collective regularizations of user-user, item-item and user-item similarities. SemRec (Shi et al., 2015) highlighted the users' loved and hated items through their previous interactions history by learning the sampled path-patterns for positive and negative user preferences. Although propagation of preferences on sampled path-patterns generate accurate recommendations, explicit and lengthy tuning of hyper-parameters is a drawback of SemRec in the case of long Meta paths. The working mechanism of SemRec is shown in Fig. 4.

Similarly, FMG (Zhao et al., 2017) suggested that just users' loved items need to be considered because normally people rate their liked items. For this purpose, they tried to overwhelm the representation deficiencies of Meta-path by substituting Meta-paths with Meta-graphs.

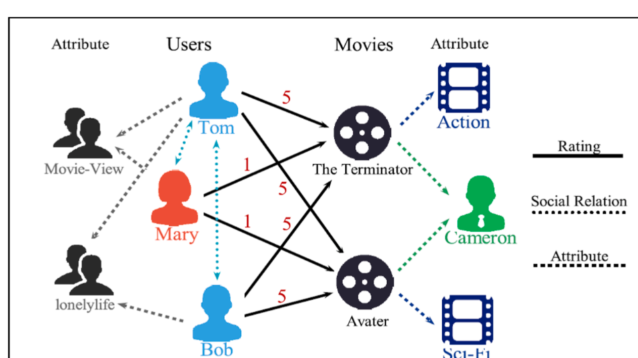

Fig. 4. Movie RS considers Entities and Relations as a Network of Heterogeneous Information Systems (Shi et al., 2015).

[42] Graph Representation Learning.

They claimed that FMG can accurately detect similarity options between user-user, item-item and user-item because Meta-graphs contain richer structural information as compared to Meta-path. FMG utilized MF to generate user-item latent vectors from each Meta-graph, and used factorization machine to combine the features of users and items to calculate the preferences. Factorization machine deals with the patterns of relations among entities; since it represents the interactions of entities across the Meta-graph. From yelp data, an exemplary food scenario of Royal House[43] is exploited in this paper, as shown in Figs. 5 and 6. Although SemRec and FMG simplified many complexities regarding Meta-paths, their strict specificities are termed as their limitations. Table 4 shows the summarization of studies about MF-based Method.

ProPPR (Catherine & Cohen, 2016) used MF for recommendation generation, and HERec (Shi et al., 2018) unified MF with the extracted Meta paths. Moreover, RuleRec (Ma et al., 2019) can be considered as a type of explainable RSs because it is comprised of explicit rules and associated rule-weights. To find the connections between entities, it mutually trained two modules, i.e. rule-learning and item-recommendation; based on a single vector by cashing the entities connection in the target graph. The working mechanism of RuleRec (i.e., rule learning and recommendation modules) and parameter's vector is described in Figs. 7 and 8 respectively.

In the target graph, the rule-learner linked the items with relevant users and converted the Meta paths to the explainable rules. Each rule is learned and assigned a specific weight. Rules and their weights are combined with users' previous interaction history by item-recommendation module via MF to generate recommendations. (Piao & Breslin, 2018) performed knowledge transferring between the recommendation and KG completion modules in item-specific domain through the proposed transfer learning approach known as CoFM[44]. For item recommendation and KG completion, FM and TransE are utilized respectively. FM, commonly applied in collaborative filtering, is a baseline framework for general latent factor model (LFM), and TransE is translation centered embedding model used to learn LFM for e_h, r, e_t to validate triplets in KG. CKE (Zhang et al., 2016a) combined item information with collaborative filtering framework to create KG embedding module. Intuitively, TransR is used to encode latent vector $\mathbf{x}_j$ of item's structural knowledge. Auto Encoder (AE) architecture is applied to extract its text $\mathbf{z}_{tex,j}$ and visual $\mathbf{z}_{vis,j}$ features; and its offset vector $\mathbf{y}_j$ is extracted from $\mathscr{M}$. Item v_j's final representation $\mathbf{v}_j$ is calculated as $\mathbf{v}_j = \mathbf{x}_j + \mathbf{y}_j + \mathbf{z}_{tex,j} + \mathbf{z}_{vis,j}$. User preference $\check{\mathscr{Y}}_{(i,j)}$ is calculated through the inner product of the concerned entities as $\mathsf{u}_i^T \mathbf{v}_j$. The user u_i's latent vector u_i is acquired in the same way. Finally, the ranking $v_{j_1} > v_{j_2} > \ldots > v_{j_n} \rightarrow \mathbf{u}_i^T v_{j_1} > \mathbf{u}_i^T v_{j_2} > \ldots > \mathbf{u}_i^T v_{j_n}$ recommended v_{j_1} to u_i. The authors claimed that the inclusion of structural knowledge to the item's information has enhanced the performance. Similarly, SHINE (Wang et al., 2018a) predicted the sentiments among entities in the KG to generate the recommendations about celebrities. It embedded sentiments $\mathscr{N}_{en}$, social $\mathscr{N}_{sn}$ and profile $\mathscr{N}_{pn}$ information networks of users and celebrities as side information through AE technique and aggregated as the ultimate representations of users and celebrities. Finally, it applied $\check{\mathscr{Y}}_{(i,j)} = f(\mathsf{u}_i, \mathbf{v}_j)$ to generate the recommendations.

4.3.2. Attention mechanism (AM)-based methods

In Deep Neural Networks, Attention Mechanism is an attempt to mimic the actions of human brain of selectively focusing on few appropriate options and ignoring the rest. Various researchers exploited AM techniques to enhance the performance of KG-based recommendation approaches. For instance, RippleNet (Wang et al., 2019b) – pioneer in presenting the idea of preference propagation – performed entity embedding and sampled the ripple sets. It used ripple sets to refine the user representations as:

$$p_i = \frac{\exp\left(\mathbf{v}_j^T \boldsymbol{R}_i \boldsymbol{e}_{h_i}\right)}{\sum_{\left(e_{h_k}, r_k, e_{t_k}\right) \in \mathfrak{S}^1_{u_i}} \exp\left(\mathbf{v}_j^T \boldsymbol{R}_k \boldsymbol{e}_{h_k}\right)} \tag{13}$$

Where $\boldsymbol{R}_i \in \mathbb{R}^{k \times d}$ shows the embedding of relation $\boldsymbol{r}$, and $\boldsymbol{e}_{h_i} \in \mathbb{R}^d$ is e_h embedding as $\mathfrak{S}^k_{u_i}$.

The relation $\boldsymbol{r}$ is used to calculate the similarity between e_h and v_j. Therefore, the order-1 response of user's past interactions is calculated through the following equation:

$$O^1_{u_i} = \sum_{\left(e_{h_i}, r_i, e_{t_i}\right) \in \mathfrak{S}^1_{u_i}} p_i \boldsymbol{e}_{t_i} \tag{14}$$

Where $\boldsymbol{e}_{t_i} \in \mathbb{R}^d$ represents the embedding of e_t as $\mathfrak{S}^k_{u_i}$.

RippleNet adopted two variations of ripple sets, i.e., information Propagation and Aggregation with respect to the application of Attention Mechanism. In information propagation, tail-entities are assigned average weights of resemblance based on their similarities with other tail-entities, their concerned heads and other associated entities. While in information aggregation, it used AM to biasedly average the representations of neighborhood with respect to a given entity. RippleNet applied multilevel AM to acquire higher-order users' preferences in KG. The general representation of users' order-h response is shown via $O^h_{u_i}$, where $h = 2, 3, \cdots, L$. Thereby, users' final representation u_i is sum of $O^1_{u_i}$ and $O^h_{u_i}$, such that, $\mathsf{u}_i = O^1_{u_i} + O^2_{u_i} + \cdots + O^L_{u_i}$. The preferences are propagated in KG in this way. Conclusively, recommendations are generated with user preferences that are calculated via $\check{\mathscr{Y}}_{(i,j)} = \sigma(\mathsf{u}_i{}^T \mathbf{v}_j)$, where $\sigma()$ is the sigmoid function. Table 5 shows the summarization of studies about AM-based Methods.

AKUPM (Tang et al., 2019) modeled the users with respect to their previous click history. It learned the representations of users and items with TransR and the relations among them with a self-AM, and propagated the acquired information over KG with bias. After that, AKUPM combined the embeddings of various hop-neighbors of selected items via self-AM network to achieve the final representations of users. The working flowchart of AKUPM is shown in Fig. 9, whereas it's self-attentive mechanism is presented in Fig. 10. The AM-network is used to keep the track of entity propagations in KG.

Recently, RCoLM (Li et al., 2019c), an extension of AKUPM, considered AKUPM as a backbone of the approach and jointly learned the modules of recommendation and KG completion. RCF (Xin et al., 2019) embedded items' relation-type and relation-value via its hierarchical description. To maintain v_j - v_k relational structure, RCF used DistMult to augment KGE module and utilized AM to disjointedly model the type and value level preferences of users. Moreover, RCF acquired better performance from the combined training of *KG-relation* and *recommendation* modules. Similarly, MKR (Wang et al., 2019d) included the execution of two modules, i.e., KGE and recommendation. It learned the representations of those entities that are connected with the items through KGE and trained the latent representations of u_i and v_j through the recommendation module.

4.3.3. Deep learning (DL)-based methods

Deep Learning – a type of Machine Learning related to algorithms enthused by the functions and structure of brain i.e., *artificial neural networks – adopts multiple layers of computation to extract gradually higher-order relations from the KG data.* DL-enhanced methods are highly applied by the recent researchers to improve the performance of KG-based recommendation processes. For instance, MCRec (Hu et al., 2018) used AM with MF and Convolutional Neural Network (CNN) to describe the context of interactions of user-item pairs through learning explicit representations of Meta paths. After the pair-wise embedding of users and items through lookup layer, MCRec defined $\mathscr{D}$-meta paths to connect entities; and modeled $\mathscr{K}$-path occurrences of each $\wp$ in $\mathscr{D}$ where $\wp \in \mathscr{D}$. CNN is used to embed the path occurrences to obtain the

[43] https://www.yelp.com/biz/royal-house-new-orleans.

[44] Co-Factorization Model.

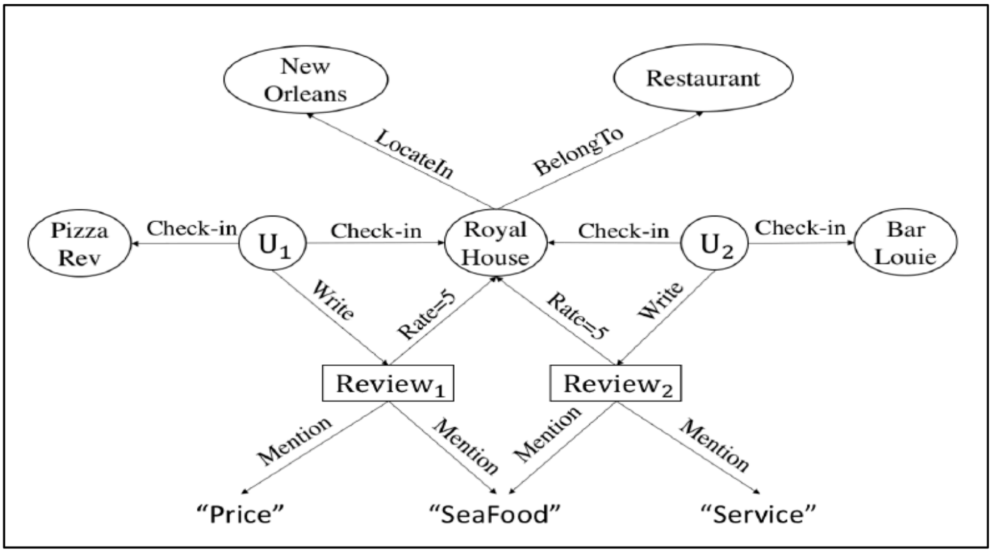

Fig. 5. Heterogeneous Information Network – Yelp Dataset-based Web of Royal-House (Zhao et al., 2017).

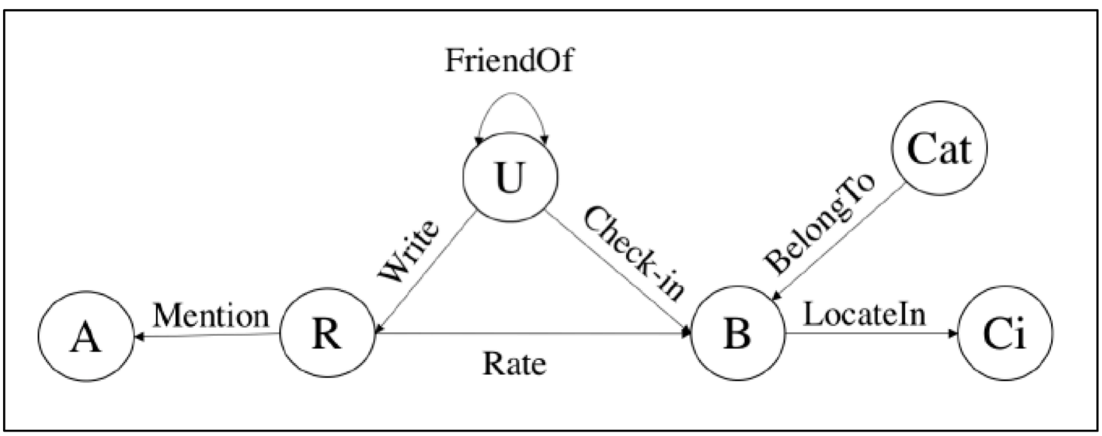

Fig. 6. Schema of Exploitation of Heterogeneous Information Network for Recommendation; Abbreviations used – A: Aspects Mined from Reviews; B: Business; Ci: City; U: Users; R: Reviews; Cat: Item-Category (Zhao et al., 2017).

embedded representations of each path occurrence (i.e., $\mathcal{E}_p$) and max-pooling operation is applied on all $\mathcal{E}_p$ to get the Meta-path embedding; represented through $\mathcal{E}_{\wp}$. User-item representations are modified and adjusted, and all $\mathcal{E}_{\wp}$ are unified to get UIE (i.e., $\mathcal{E}_I$) through AM. Finally, $\check{\mathcal{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathcal{E}_I)$ is used to calculate the preference scores.

KPRN (Wang et al., 2019g) applied AM with Recurrent Neural Network (RNN) to acquire entity and relation embedding to build user-item path-sequence; applied LSTM to encode it, and used $\check{\mathcal{Y}}_{(i,j)} = f(\mathcal{E})$ from $\check{\mathcal{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathcal{E}_I)$ to obtain user to item preferences in each path. Through the weighted pooling layer, prediction scores of all independent paths in the concerned path-sequence are unified into the ultimate preference score that infer recommendations. Moreover, without any prior definitions of Meta-path, RKGE (Sun et al., 2018b) automatically mined users, items and their relations through path-sequences. With the semantic relations in the graph, first RKGE followed the Meta path $\wp(u_i, v_j)$ that connected user u_i and item v_j under the applied constraints (i.e., the length of path-sequence, etc.). Second, it encoded the entire user-item path-sequence through a designed recurrent network. The paths are interlinked on user-item embedding sequences. Ultimate hidden states (i.e., $\mathcal{E}_h$) about the encoded paths are obtained from $\mathcal{E} = g(\mathcal{E}_p)$, and average pooling operation is applied to unify the set of concerned

Table 4
MF and CF-based Methods: Information about the Methods, Venues, Applied Frameworks, Used Datasets, Evaluation Parameters and Publication-details of MF-based methods. Abbreviations used – AP: Amazon Products, DM: Douban-Movie, DB: Douban-Book, YC: Yelp Challenge, CY: CIKM-Yelp, CD: CIKM-Douban, ML: Movie Lens, DBLP: Digital-Bibliography and Library-Project, MAE: Mean Absolute Error, RMSE: Root Mean Square Error, NDCG: Normalized Discounted Cumulative Gain, Prec: Precision.

Method	Venue	Framework	Datasets	Evaluation Parameters	References
RuleRec	WWW	MF	AP	(Recall, NDCG, MRR)@K	(Ma et al., 2019)
HERec	TKDE	MF	DM, DB, YC	MAE, RMSE	(Shi et al., 2018)
FMG	KDD	MF	YC, AP, CY, CD	RMSE	(Zhao et al., 2017)
ProPPR	RecSys	MF	YC, ML-100 K	Prec@K, MRR	(Catherine & Cohen, 2016)
SemRec	CIKM	MF	DM, YC	RMSE, MAE	(Shi et al., 2015)
Hete-CF	ICDM	MF	DBLP	MAE, RMSE	(Luo et al., 2014)
HeteRec-p	WSDM	MF	ML-100 K, YC	Prec@K, MRR	(Yu et al., 2014)
Hete-MF	IJCAI	MF	ML-100 K	MAE, RMSE	(Yu et al., 2013a)
HeteRec	RecSys	MF	ML-100 K, YC	Prec@K, MRR	(Yu et al., 2013b)

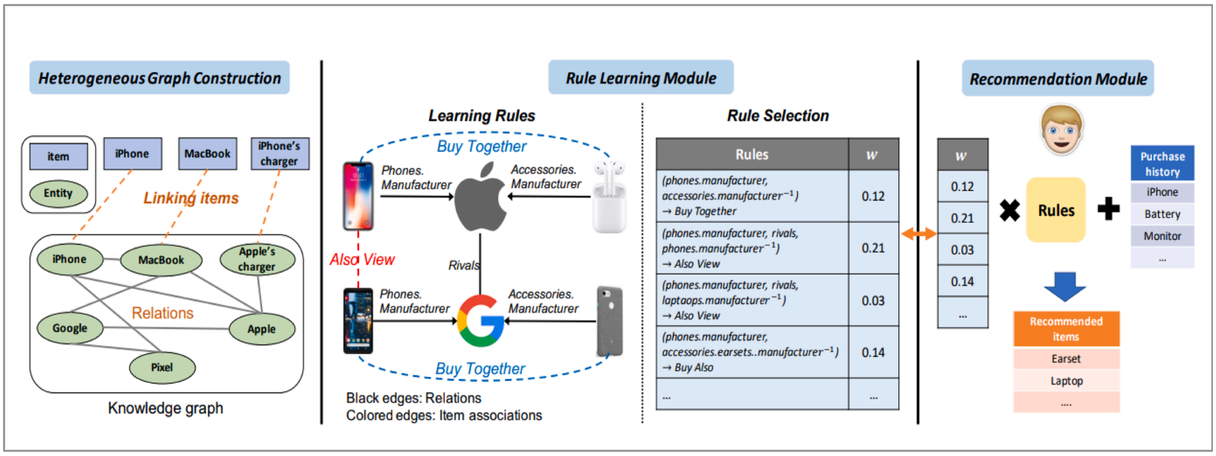

Fig. 7. The working mechanism of RuleRec Model – an overview in parts; (a) constructed a heterogeneous information graph (i.e., a local subgraph), (b) rule-learning module is learned, (c) recommendation-module is learned, (d) part (b) & (c) share the same parameter's vector **w** for synchronization (Ma et al., 2019).

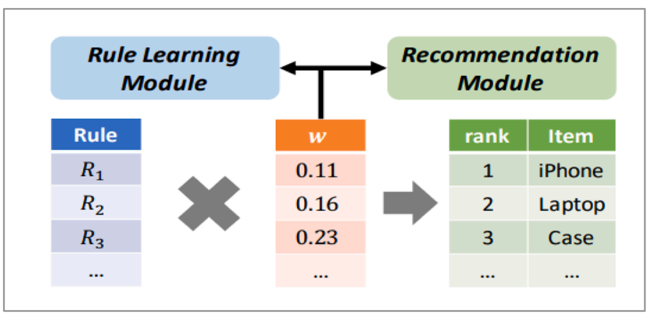

Fig. 8. The overview of RuleRec Model's multi-task learning (i.e., joint execution of Rule-learning and Recommendation modules) approach through sharing the same parameters' vector **w**. It is an example of Matrix Factorization actually (Ma et al., 2019).

$\mathscr{E}_h$ to model the semantic relation $\mathscr{E}$ between user u_i and item v_j. Therefore, equation $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathscr{E}_I)$ is modified to $\check{\mathscr{Y}}_{(i,j)} = f(\mathscr{E})$, where $\mathscr{E}$ is used to calculate the user u_i's preference towards item v_j. We explain the working mechanism of RKGE model with the help of following scenario as well as via Fig. 11. For instance, consider two *meta-paths*, i.e., "*Bob* $\xrightarrow{rate}$ *TT* $\xrightarrow{categorized-by}$ *Drama* $\xrightarrow{categorize}$ *SPR*" and "*Bob* $\xrightarrow{rate}$ *TT* $\xrightarrow{directed-by}$ *Steven Spielberg* $\xrightarrow{direct}$ *SPR*", where first relates to the "same genre" and second is about the "same director" whose movies have already been watched by the Bob. It is hereby inferred that Bob's new preferences either base on previous genre or director. Therefore, the model can recommend a drama from the same genre (i.e., GWH) or a movie directed by Steven Spielberg (i.e., SL) to the Bob. It is hereby concluded that different relations connecting the same nodes often portray different semantics with respect to the paths.

EIUM (Huang et al., 2019a) used $\mathscr{E} = g(\mathscr{E}_{\mathscr{P}})$ and $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j, \mathscr{E}_I)$ for sequential recommendation by collecting the dynamic interests of users with implementation of AM via CNN. EIUM encoded and unified such paths that connect u_i-v_j pairs to learn their paired interaction-embedding $\mathscr{E}_I$. Further, it applied AM on interactions' sequence to obtain the dynamic preference-embedding $\mathscr{E}_{\mathscr{P}}$. Thus, user to item preference calculation model is defined as $\check{\mathscr{Y}}_{(i,j)} = f(\mathscr{E}_I, \mathscr{E}_{\mathscr{P}})$. DKN (Wang et al., 2018b) – the framework is shown in Fig. 12 – is influenced from Kim-CNN (Zhang & Wallace, 2015). DKN used AM with CNN to aggregate the text embedding and applied TransD to acquire knowledge

Table 5
Attention Mechanism (AM): Information about the Methods, Venues, Applied Frameworks, Used Datasets, Evaluation Parameters and Publication-details of AM-based methods. Abbreviations used – ML: Movie Lens, KB: KKBox, LF: Last FM, BC: Book Crossing, BN: Bing News, P&R: Precision and Recall, MR: Mean Rank, HR: Hit Ratio, NDCG: Normalized Discounted Cumulative Gain, F_1: F Measure/F_1 Score, MF_1: Micro-F_1, Acc: Accuracy, AUC: Area under the Curve.

Method	Venue	Framework	Datasets	Evaluation	References
MKR	WWW	AM	ML-1 M, BC, LF, BN	AUC, Acc, P&R@K	(Wang et al., 2019d)
RCF	SIGIR	AM	ML-100 K, KB	HR, MRR, NDCG	(Xin et al., 2019)
RCoLM	IEEE Access	AM	ML-1 M, BC	Acc, AUC, (NDCG, P&R, F_1)@K, MR, (HR)@N	(Li et al., 2019c)
AKUPM	KDD	AM	ML-1 M, BC	Acc, AUC, (P&R, F_1)@K	(Tang et al., 2019)
RippleNet	CIKM	AM	ML-1 M, BC, BN	Acc, AUC, (P&R, F_1)@K	(Wang et al., 2019b)

level embedding of entities $\mathbf{v}_j$ to recommend news v_j to the end users, where $v_j \in \mathbf{v}_j$. User representation u_i is obtained from the unification of previously interacted news embedding $\mathbf{v}_I = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_N\}_i$ through AM to find the users' current interest about the potential news. For each instance in the previously interacted news instances (e.g., $v_k|k = (1, 2, \cdots, N)_i$), the attention weight and u_i's final embedding u_i is calculated through the following equations respectively:

$$\alpha_{(v_k, v_j)} = \frac{\exp(g(\mathbf{v}_k, \mathbf{v}_j))}{\sum_{k=1}^{N_i} \exp(g(\mathbf{v}_k, \mathbf{v}_j))} \tag{15}$$

Where $\alpha_{(v_k, v_j)}$ is the attention weight and $g()$ is used to work as a Deep Neural Network (DNN) layer.

$$\mathrm{u}_i = \sum_{k=1}^{N_i} \alpha_{(v_k, v_j)} \mathbf{v}_k \tag{16}$$

Finally, $\check{\mathscr{Y}}_{(i,j)} = f(\mathrm{u}_i, \mathbf{v}_j)$ is used to calculate the preferences to generate recommendations.

KSR (Huang et al., 2018) applied AM with RNN and formulated user

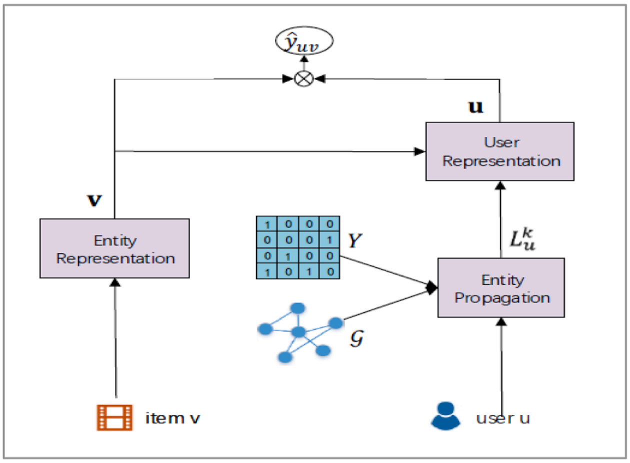

Fig. 9. General working flowchart of the AKUPM framework (Tang et al., 2019).

preferences from sequential interactions to generate sequential recommendations through unification of GRU and KKM[45] networks. User's sequential interaction based preferences $\mathscr{V}_i^t$ are acquired via GRU and attribute based preferences a_i^t are modeled via KKM with TransE learned knowledge base. Therefore, at time interval t, user and item latent vectors are $\mathbf{u}_i^t = \mathscr{V}_i^t \oplus a_i^t$ and $v_j^t = \mathbf{u}_i^t \odot \left(\beta_j \odot \gamma_j\right)$ respectively, where $\odot$ is the dot product, $\oplus$ is concatenation operation, and β_j is GRU-based embedding and γ_j is KG-based item embedding. Latent vectors u_i^t and $\mathbf{v}_j^t$ are transformed and user preferences are ranked for recommendations with respect to the results obtained from $\check{\mathscr{Y}}_{(i,j)} = f(u_i, \mathbf{v}_j)$. CGAT (Yang et al., 2020) applied AM with RNN and Graph Neural Network (GNN) and exploited entity's local and non-local context information (CI) in the KG for recommendation. CGAT acquired the local CI through user-based graph-AM with personalized user preferences and non-local CI through RNN with biased random walk sampling technique. The general framework of CGAT is displayed in Fig. 13.

Many researchers applied user-item graph based preference propagation for recommendation via AM and GNN. For example, KGCN (Wang et al., 2019e) applied AM and GNN, and aggregated the embeddings from L-hop neighbors of the interacted item to itself to achieve

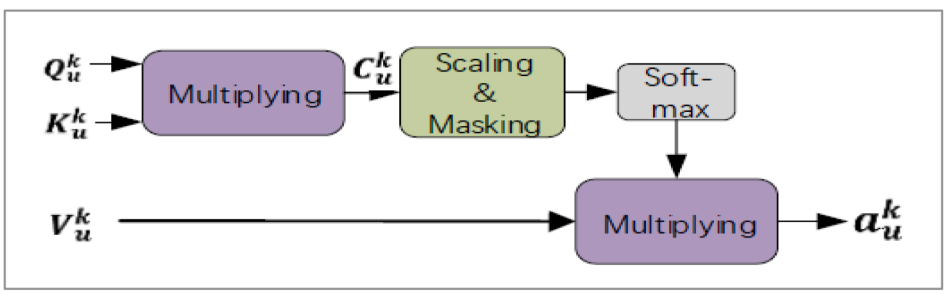

Fig. 10. Specific working flowchart of self-attention mechanism of AKUPM (Tang et al., 2019).

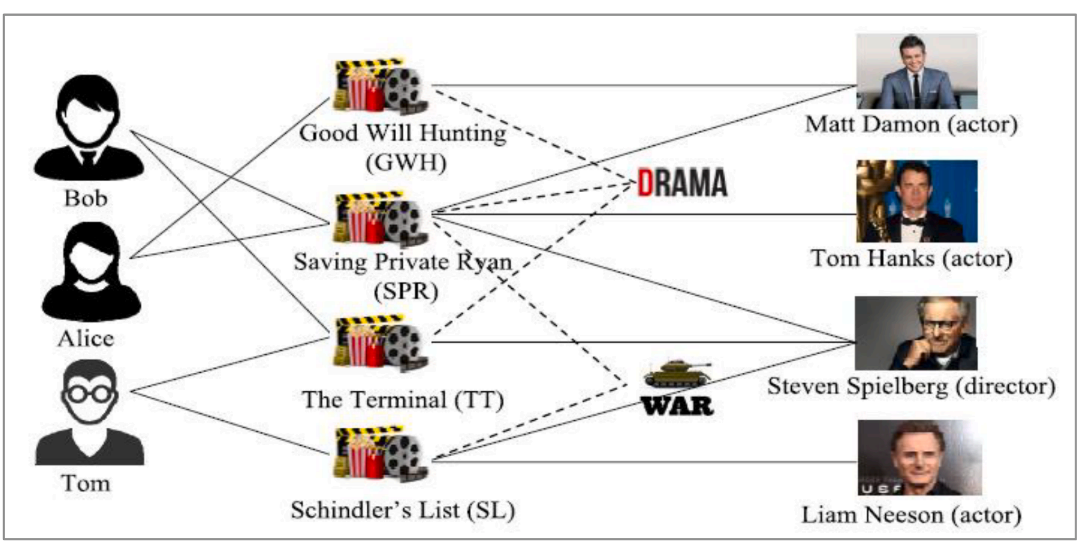

Fig. 11. A knowledge graph from movies' dataset having customers (users), movies, starrers, genres and directors as entities; and rating, categorizing, starring and directing as relations among entities (sun et al., 2018b).

[45] Gated Recurrent Unit (GRU) and Knowledge-enhanced Key-value Memory (KKM).

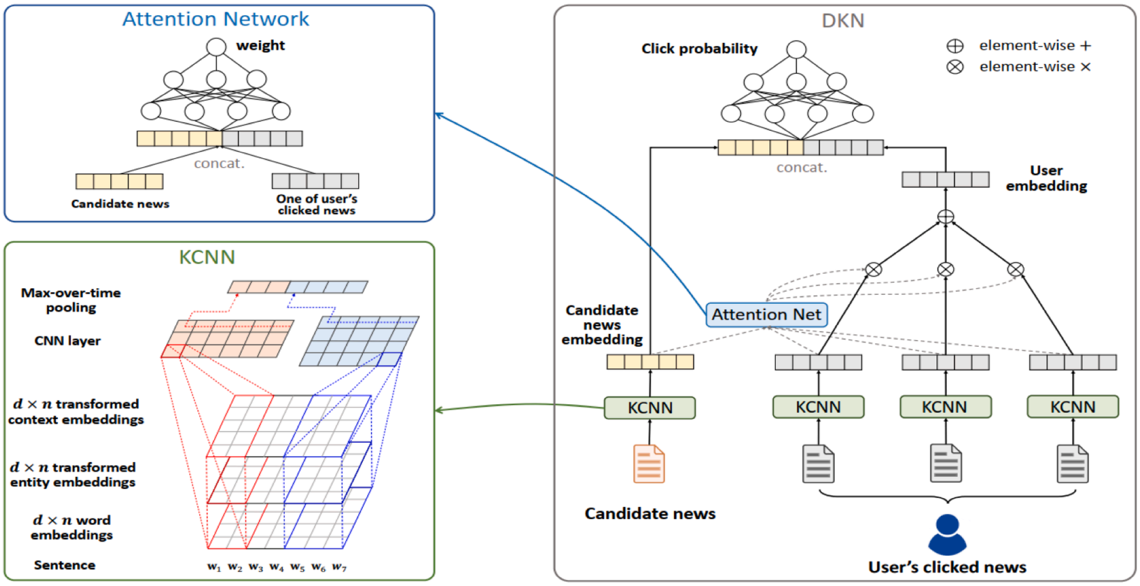

Fig. 12. The DKN Framework (Wang et al., 2018b).

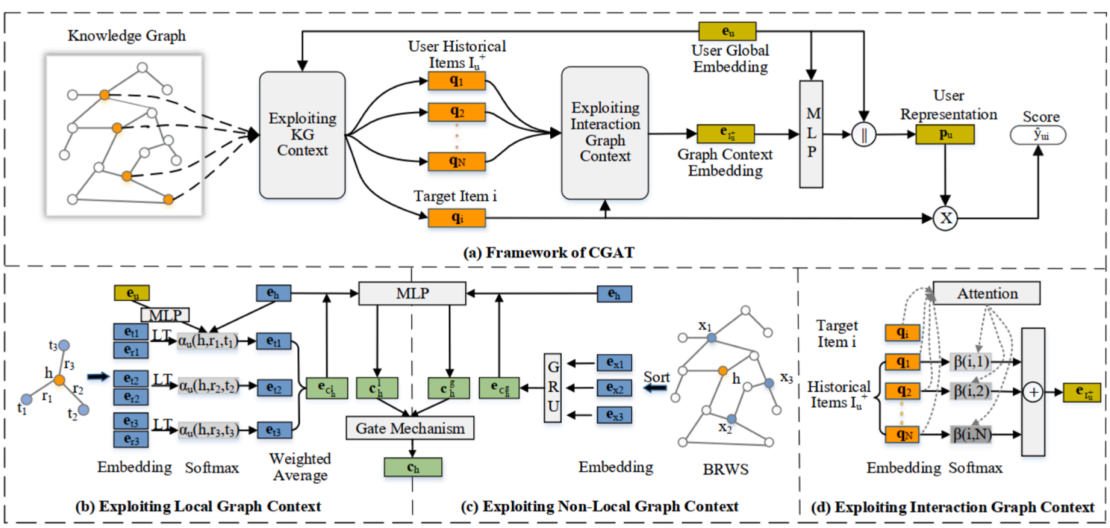

Fig. 13. The CGAT Framework. (a) Upper part from left to right, uses knowledge graph context and user-interactions' graph context to calculate the preferences of user for a potential item; (b) Uses local KG context by exploiting specific user-based graph AM on KG-data; (c) Using non-local KG context via biased random walk based sampling technique and GRU module; (d) Using the context of user-interactions' graph by exploiting item-based AM on users' previous interacted items (Yang et al., 2020).

the final representations of the candidate item e_j. First, KGCN declared e_j's neighbors and then recursively declared the neighbors of e_j's neighbors and so on. Lastly, it declared a fixed number of neighbors of all of the entities in KG. KGCN started from the upper layer (i.e.,L-hop neighbor), and iteratively replaced the values of k in the descending order from Lto1 (i.e.,$L, L-1, \cdots, 1$) in $e_h = \mathrm{Ag}(e_h, e_t)$, where $t = \mathfrak{S}^k_{e_j}$ for all of the k- hop neighbors to achieve the final representations of the internal entities. Multi-hop neighbor's information is internally propagated to e_j via inward propagation – a KG process of features propagation. The final representation of e_j is comprised of e_j's initial representation and its multi-hop neighbors' information.

KGCN learned the items' representations from the embeddings of items' far-neighbors through inward propagation, while RippleNet learned the preferences from users' previous interactions through outward propagation. Recently, KGNN-LS (Wang et al., 2019c) combined AM and GNN to enhance KGCN through the integration of label smoothness (LS) technique in the existing framework. First, it modelled the learning process to acquire the e_j's inclusive representations through users' interaction information. Then, it propagated the labels of users' interactions on KG to collect users' preferences. Intent-GC (Zhao et al., 2019b) exploited those rich behaviors in the KG that are deemed as closely related to the users' interests to generate rich recommendations. They claimed that they have designed a faster GNN to support the scalability in the framework.

RippleNet and its descendants mainly focused on item-based embedding propagation. For instance, KGAT (Wang et al., 2019f) acquired higher order user-item relations with embedded information propagation. First, it acquired the initial representations of entities through TransR, and then it initiated the propagation based on the corresponding entities. In this way, the multi-hop neighbors are

Table 6
Deep Learning-based Methods: Information about the Methods, Venues, Applied Frameworks, Used Datasets, Evaluation Parameters and Publication-details of DL-based methods. Abbreviation used – AM: Attention Mechanism, GNN: Graph Neural Network, BC: Book-Crossing, BN: Bing-News, ML: Movie-Lens, LF: Last-FM, AB: Amazon-Book, YC: Yelp-Challenge, DF: Dianping-Food, AT: Alibaba's Taobao, AP: Amazon Products, AUC: Area Under the Curve, MRR: Mean Reciprocal Rank, Acc: Accuracy, F_1: F Measure/F_1 Score, Prec: Precision, P&R: Precision and Recall, NDCG: Normalized Discounted Cumulative Gain, HR: Hit-Ratio, AS: Applied Sciences, IS: Information Sciences.

Method	Venue	Framework	Datasets	Evaluation	References
CGAT	arXiv	GNN, RNN, AM	LF, ML-1 M, BC	P&R@K, HR@K	(Yang et al., 2020)
KGPolicy	WWW	GNN, AM, MF, RL	AB, LF, YP	Recall@K, NDCG@K	(Wang et al., 2020b)
DKN	WWW	CNN, AM	BN	F_1, AUC	(Wang et al., 2018b)
KSR	SIGIR	RNN, AM	LF, ML-1&20 M, AB	MAP, MRR, HR, NDCG	(Huang et al., 2018)
RippleNet	TOIS	GNN, AM	ML-20 M, BC, BN	AUC, Acc, (P&R, F_1)@K	(Wang et al., 2019b)
KGCN	www	GNN, AM	ML-20 M, BC, LF	AUC, F_1, Recall@K	(Wang et al., 2019e)
KGAT	KDD	GNN, AM	AB, LF, YC	(Recall, NDCG)@K	(Wang et al., 2019f)
KGCN-LS	KDD	GNN, AM	ML-20 M, BC, LF, DF	Recall@K, AUC	(Wang et al., 2019c)
KNI	KDD	GNN, AM	BC, ML-1,20 M, AB	AUC, Acc, (P&R, F_1)@K	(Qu et al., 2019)
AKGE	arXiv	GNN, AM	ML-1 M, LF, YP	(Hit, NDCG)@N	(Sha et al., 2019)
KPRN	AAAI	RNN, AM	ML-1 M, KB	(hit, ndcg)@K	(Wang et al., 2019g)
RKGE	RecSys	RNN, AM	ML-1 M, YC	Prec@N, MRR	(Sun et al., 2018b)
EIUM	MM	RNN, AM	ML-20 M	AUC, MAP, (Hit, NDCG)@N	(Huang et al., 2019a)
MCRec	KDD	CNN, AM, MF	ML-100 K, LF, YC	(Prec, Recall, NDCG)@K	(Hu et al., 2018)
Intent-GC	KDD	GNN	AT, AP	AUC, MRR	(Zhao et al., 2019b)
NACF	AS	CNN, CF	ML-20 M, LF, DF	(Prec, Recall, NDCG)@K	(Zhang et al., 2020a)
DKEN	IS	DNN	ML-1 M, BC, LF	AUC, ACC	(Guo et al., 2020)

iteratively interacted by the information of the entity from which the propagation was initiated. The process is defined as $e_i^{k+1} = \mathrm{Ag}(e_i^k, e_{\mathfrak{S}_{e_i}^{k+1}})$, where $k = 0,1,\cdots,L-1$, e_i^k is k-hop neighbor's connection, and e_i^k shows the initial representations of the entities if $k = 0$. To get the final representations e_i^{φ}, all e_i^k embeddings of user's and item's representations are improved to final users $e_{\mathscr{u}}^{\varphi}$ and items $e_{\mathscr{v}}^{\varphi}$ representations respectively, with respect to their corresponding neighbors and unified in e_i^{φ} with bias. Table 6 represents the summarization of studies about the Attention Mechanism with NN techniques. Finally, the user preferences are calculated through $\mathscr{Y}_{(\mathscr{u},\mathscr{v})} = e_{\mathscr{u}}^{\varphi T} e_{\mathscr{v}}^{\varphi}$ to generate the recommendations. RippleNet (Wang et al., 2019b) aggregated the outward and inward propagation of user preferences to address the limitations faced by the individual techniques (i.e., path or embedding based method). Moreover, KNI (Qu et al., 2019) proposed their approach to jointly refine the embeddings of users and items. For this purpose, KNI took into account the historical interactions happened among the neighbors of users and items in the graph.

AKGE (Sha et al., 2019) propagated user-item paired information in the graph to learn the representations of $\mathscr{u}_i$ and $\mathscr{v}_j$. First, AKGE used TransR to learn entity embeddings in KG, and then created a $\mathscr{u}_i$ -$\mathscr{v}_j$ relation-based subgraph $\mathscr{S}_{KG}$ through modelling $\mathscr{u}_i$ -$\mathscr{v}_j$ inter-connected paths based on pair-wise distance between them (i.e., the paths). During the construction of subgraph, AKGE focused on the idea *"lesser is the distance among entities in the Euclidian space, more similar are the entities with each other; and* vice versa". This way, $\mathscr{S}_{KG}$ acquired higher order $\mathscr{u}_i$ -$\mathscr{v}_j$ relevant-pairs by filtering out the less-appropriate entities from the KG. To achieve the final preferences about $\mathscr{u}_i$ -$\mathscr{v}_j$ pairs in the $\mathscr{S}_{KG}$, AKGE applied AM-based GNN model to propagate the information of entities and their neighbors on the subgraph. AKGE supports and generates explainable recommendations; and the working mechanism of explainability of the recommendations provided by AKGE is described in Fig. 14.

NACF (Zhang et al., 2020a) identified the potential interests of users based on information collected from the neighbors of the users via propagation over KG. In other words, it collected the neighborhood information of the entities to mine the potential relations among users and items to guess users' possible preferences about the unseen items of their interest. It is an AM and CF-based joint approach. DKEN (Guo et al., 2020) exploited the impact of data-exchange (through CIS[46] layer) between the implicit and explicit semantics of information based on user-to-item interactions and KG-features respectively. Additionally, the proposed model performed data-exchange to acquire a better grip on the semantical and hierarchical structure of information-flow in the KG. Moreover, it is a DL-based approach that performed information propagation in the KG; and applied *hottest* and *k-largest* node samplers as the information aggregators. Table 7 summarizes the specifications of the rest of the KG-based recommendation approaches.

4.3.4. Reinforcement learning (RL)-based methods

RL – a subtype of Machine Learning and Artificial Intelligence – is dynamic programming that exploits the systematic techniques of reward and penalty to train the concerned algorithms. RL attempts to optimize the performance with appropriate actions by maximizing the reward. It is also utilized in KG-based recommendation to improve the performance. For instance, PGPR (Xian et al., 2019) proposed RL mechanism to find the realistic paths between $\mathscr{u}_i$-$\mathscr{v}_j$ through designing the concerned problem according to Markov Decision Process. In this process, they designed an algorithm to search $\mathscr{u}_i$-$\mathscr{v}_j$ path, defined path to path transition strategies, set conditions to terminate the paths, and formulated how to learn the model to sample $\mathscr{u}_i$-$\mathscr{v}_j$ paths through a trained agent. The working mechanism of the proposed approach is represented in Fig. 15. Similarly, EKar (Song et al., 2019a) used RL approach to provide item's recommendations to the users. It formulated the task as a chronological decision process, where entities are considered as states and walks among entities as actions to acquire the meaningful connections among users and items for recommendation. It designed the scoring function according to the current relevant approaches and trained a policy-guided function with the help of policy gradient techniques.

KGPolicy (Wang et al., 2020b) used RL with AM, MF and GNN to discover high quality negatives (i.e., missing data, negative samples, false signals, etc.) in the KG to control the negative sampling. KGPolicy initiated to learn the agent over positive interactions of the users with the targets, progressively obtained knowledge-aware missing data, false samples (negative signals), and finally generated a potential negative entity (i.e., missing item to learn the recommendation technique). It embedded users $\mathscr{u}$ and items $\mathscr{v}$ information to the vector space; and

[46] Cross Information Share.

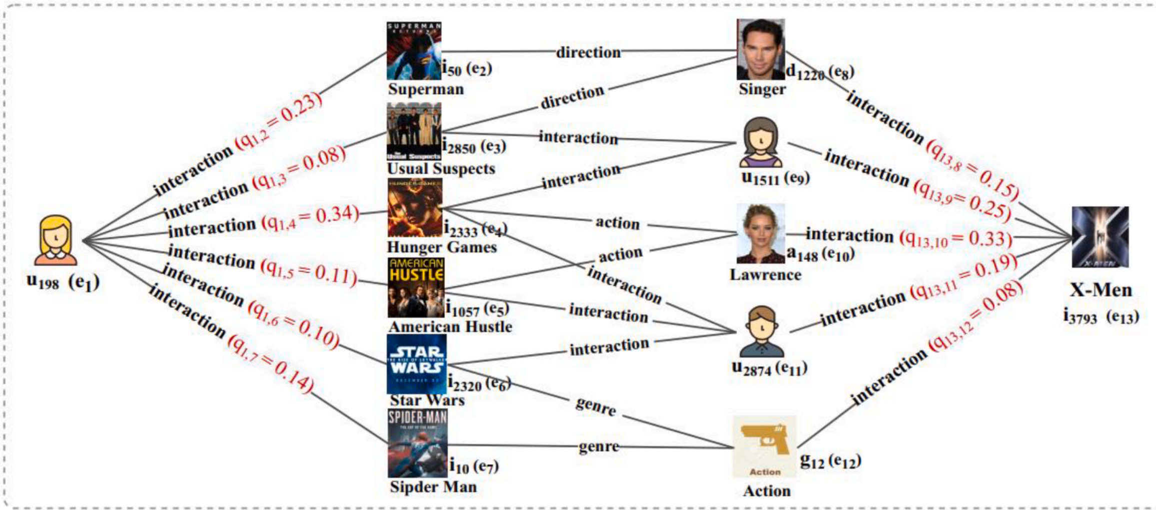

Fig. 14. A descriptive prototypical application scenario-based on a toy example acquired from the MovieLens-1 M dataset to designate explainability of the recommendations provided by AKGE approach. The self-constructed local subgraph that inter-links the pairs of users and items, for instance (u_{198} – X-Men), is comprised of movies (e.g., Superman); users (e.g., u_{198}); actors (e.g., Lawrence); directors (e.g., Singer); and genres (e.g., Action) as entities; while direction (e.g., Superman, that is directed by Singer); interaction (e.g., u_{198} interacts with the Superman); action (e.g., Hunger Game is starred by Lawrence) and genre (Star Wars goes to the action) as relations among the mentioned entities. The entities $e_1, \cdots, e_{13}$ represent the mapped entity-indices in the subgraph and the values in red (e.g., $q_{1,2} = 0.23$), that are attached on the link, show the attention-weights of neighbors (e.g., e_2) w.r.t target entity (e.g., e_1). For further details and thorough contemplation, please consult the case study of research question 2 in (Sha et al., 2019).

Table 7
The rest of the Approaches. Information about the Methods, Venues, Applied Frameworks, Used Datasets, Evaluation Parameters and Publication-details. Abbreviations used – FARF: Fairness Aware Ranking Framework, AM: Attention Mechanism, AE: Auto Encoder, BF: Bayesian Framework, GAN: Generative and Adversarial Network, RL: Reinforcement Learning, FM: Factorization Machines, ML: Movie-Lens, LF: Last-FM, AT: Alibaba's Taobao, Db14: DB-books2014, SW: Sina-Weibo, IB: Intent-Books, LT: LibraryThing, AP: Amazon-Products, HRR: Hit Recall Rate, F_1: F Measure/F_1 Score, MF_1: Micro-F_1, AUC: Area Under the Curve, Acc: Accuracy, Prec: Precision, Avg: Average, P&R: Precision and Recall, Ps: Pearson, Sm: Spearman, MRR: Mean Reciprocal Rank, NDCG: Normalized Discounted Cumulative Gain, HR: Hit-Ratio, MAP: Mean Average Precision, FND: Framework Not Defined.

Method	Name	Framework	Datasets	Evaluation Parameters	References
FAER	SIGIR	FARF	AP	NDCG, F_1	(Fu et al., 2020)
PGPR	SIGIR	RL	AP	P&R, HR, NDCG	(Xian et al., 2019)
Ekar	arXiv	RL	LF, ML-1 M, Db14	(HR, NDCG)@K	(Song et al., 2019a)
BEM	CIKM	BF	AT	HRR, Acc	(Ye et al., 2019)
CoFM	WWW	FM	ML-1 M, Db14	MRR, MAP, (P, R, NDCG)@N	(Piao & Breslin, 2018)
SHINE	WSDM	AE	SW	Acc, MF_1, P&R@K	(Wang et al., 2018a)
CKE	KDD	AE	ML-1 M, IB	MAP@K, Recall@K	(Zhang et al., 2016a)
KTGAN	ICDM	GAN	DM	Prec, Avg-Prec, NDCG	(Yang et al., 2018)
entity2rec	ExpSys	FND	ML-1 M, LF, LT	(P, R, SER, NOV)@N	(Palumbo et al., 2020)
KTUP	WWW	FND	ML-1 M, Db14	P&R, F_1, HR, NDCG, MR	(Cao et al., 2019)
DKFM	WWW	FND	CEM	HR, MRR	(Dadoun et al., 2019)
SED	WWW	FND	CNRec	F_1, P&R, Ps, Sm	(Joseph & Jiang, 2019)
node2vec	WWW	FND	ML-1 M	(P, R)@N, MAP, NDCG, MRR	(Palumbo et al., 2018)
ECFKG	Algorithms	FND	AP	P&R, HR, NDCG	(Ai et al., 2018)
CFKG	SIGIR	FND	AP	P&R, HR, NDCG	(Zhang et al., 2018c)
entity2rec	RecSys	FND	ML-1 M	P@N, MAP	(Palumbo et al., 2017)

applied MF to perform inner product of the embeddings to predict how likely u can choose v. To get the preferences, $\breve{\mathcal{Y}}_{uv} = f_R(u, v) = r_u^\top r_v$ is applied, where $f_R()$ is the interaction evaluation function, and r_u& r_v denote the embeddings of users u and items v respectively.

4.3.5. KG-based General recommendation methods

In this section, we discuss such approaches that either have not explicitly declared their applied frameworks or performed multi-task learning (i.e., KG completion and recommendation) in their studies. Though KGE-learned less-stable latent-vectors can be directly exploited for recommendation, the learned representations are usually refined to improve the performance. BEM (Ye et al., 2019) applied Bayesian Probabilistic Network to exploit content and behavior graphs to deal with the entities (i.e., items) for recommendation. The content graph contained the features of items such as brand, type, category, etc., and the behavior graph covered the interactions of users with-items such as view, purchase, rate, etc. BEM applied TransE on content and GNN on behavior graph to learn the initial embeddings; and refined both of the embeddings in a parallel way. In the behavior graph, such closest items are searched for the recommendation that belonged to the set of interacted items. SDPL[47] model (Rafailidis, 2019) introduced the concept of trust and dis-trust in the recommendation. They improved the accuracy of top-k recommendation by selecting the interactional users and their neighbors (i.e., friends and foes) through BPR[48] criterion-based multiple

[47] Social Deep Pairwise Learning.
[48] Bayesian Pairwise Ranking.

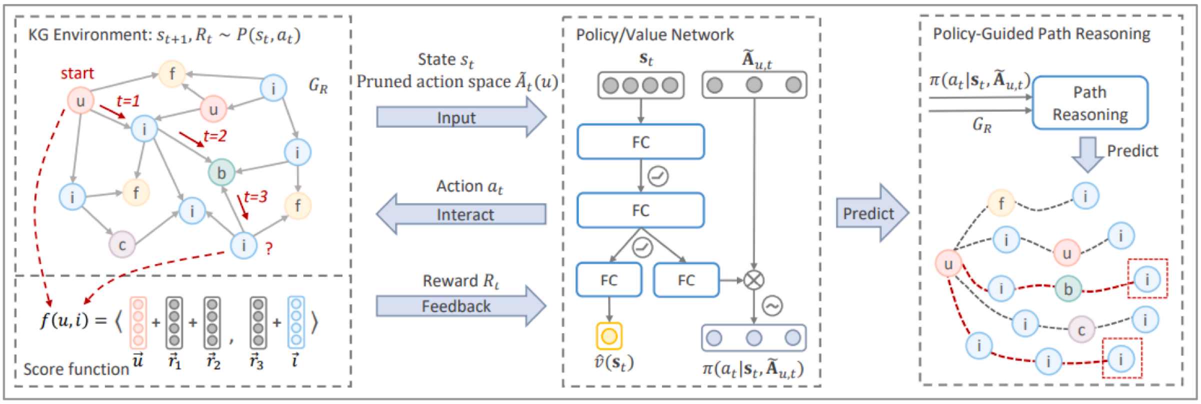

Fig. 15. The Flowchart of the working mechanism of PGPR model for recommendation. The algorithm of the proposed model targets to learn such a policy that automatically suggests potential items to the end users based on their previous interaction records. The trained algorithm is then assumed to be exploited for the path reasoning strategy to actually recommend the selected items to the users. The model is implemented via the reinforcement learning technique (Xian et al., 2019).

ranking loss function. And, (Zhang et al., 2016b) introduced a probabilistic time-based entity recommendation model. Moreover, entity2rec (Palumbo et al., 2017) and (Ai et al., 2018) incorporated the information of the users into KG and directly modeled users' preferences via creating user-item interlinked subgraphs. Similarly, CFKG (Zhang et al., 2018c) deemed user interaction or behavior with the item as a user-item relation and item features as side-information in the construction of user-item graph. The defined metric functions is:

$$\mathscr{D} = \partial(\mathrm{u}_i + \mathbf{r}_{type}, \mathbf{v}_j) \tag{17}$$

Where $\mathscr{D}$ is the result-pool that store the calculated distances, $\partial()$ is a metric function that calculates the distance between user and item with respect to a given relation, and $\mathbf{r}_{type}$ is the type of relation (e.g., $\mathbf{r}_{purchse}$). Moreover, node2vec (Palumbo et al., 2018) is designed on two aspects, such that, user-based special relation type (i.e. feedback), and item's relations with other entities in the KG. Therefore, it is clear that node2vec is a 'feedback'-based KG embedding technique to recommend items. SED (Joseph & Jiang, 2019) – a KG traversal method – provided a possible solution to the issues of cold start regarding News recommendations.

KTGAN (Yang et al., 2018) applied Metapath2Vec (Dong et al., 2017) on movies KG to learn the knowledge embedding $\mathbf{v}_j^k$ for the candidate movie $\mathscr{m}_j$, and used Word2Vec (Mikolov et al., 2013) for the movies features to learn tag embedding $\mathbf{v}_j^t$. The working mechanism is signified in Fig. 16. Average knowledge embedding of users' liked and tagged movies are represented via $\mathbf{u}_i^k$ and $\mathbf{u}_i^t$ respectively. Thus, $\mathscr{u}_i$ and $\mathscr{m}_j$'s latent vectors are defined as $u_i^{in} = u_i^k \oplus u_i^t$ and $v_j^{in} = v_j^k \oplus v_j^t$ respectively, and refined through the Generator G and the Discriminator D. Preference score functions, i.e., the Probabilistic $P_s = p_\theta(\mathscr{m}_j|\mathscr{u}_i, r)$ and the Learned $L_s = f_\phi(\mathscr{u}_i, \mathscr{m}_j)$, where r shows the extent of the relevance between user $\mathscr{u}_i$ and movie $\mathscr{m}_j$, are selected. According to the probabilistic function P_s, G suggested the movies that are relevant to the interest of user. At the system's learning, G tried to maximize the distribution of the users' truly liked movies (i.e., $d_T = p_{true}(\mathscr{m}_j|\mathscr{u}_i, r)$ through $P_s = p_\theta(\mathscr{m}_j|\mathscr{u}_i, r)$ to suggest the relevant $\mathscr{m}_j$ to $\mathscr{u}_i$). According to the learned function L_s, D distinguished the relevant pairs from the irrelevant bulk and suggested the relevant $\mathscr{u}_i$ - $\mathscr{m}_j$ connections for recommendation. Hence, the objective function is defined as:

$$\mathscr{L} = \min_\theta \max_\phi \sum_{i=1}^{M} \{\Psi[\log \mathrm{P}(\mathrm{v_j}|\mathrm{u_i})] + \Phi[\log(1 - \mathrm{P}(\mathrm{v_j}|\mathrm{u_i}))]\} \tag{18}$$

Where $\Psi = \mathrm{E}_{v_j \sim p_{true}(v_j|u_i,\ r)}, \Phi = \mathrm{E}_{\mathscr{m}_j \sim p_\theta(\mathscr{m}_j|\mathscr{u}_i, r)}$, $P(\mathscr{m}_j|\mathscr{u}_i) = \frac{1}{1+\exp(-f_\phi(\mathscr{u}_i, \mathscr{m}_j))}$ is

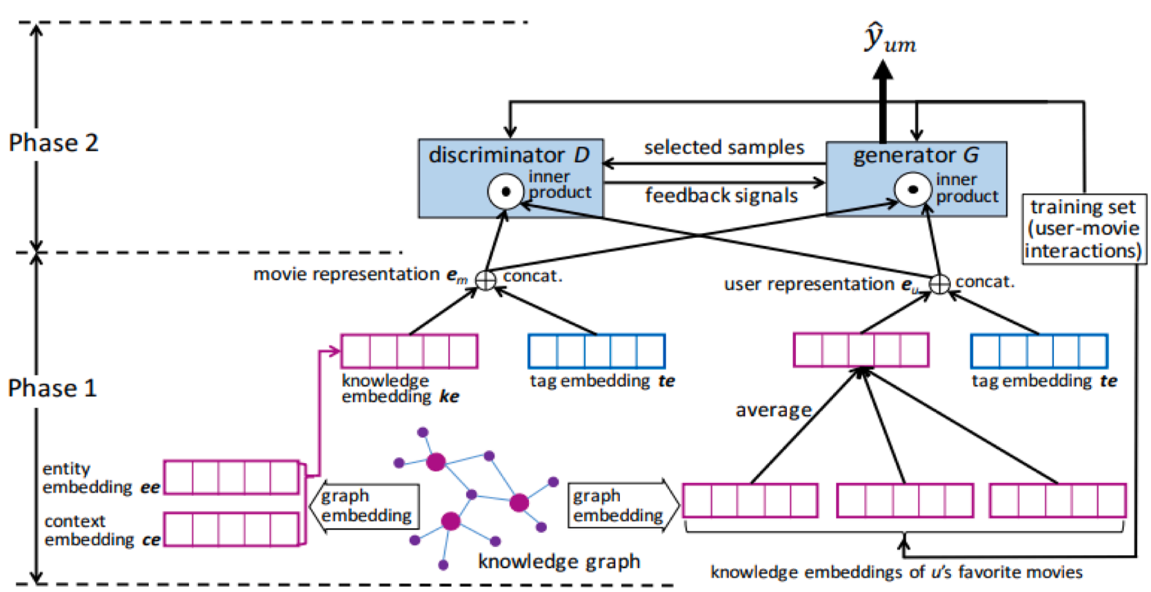

Fig. 16. KT-GAN-attributed movies recommendation framework based on movie's knowledge graph embeddings and tag embeddings from Douban dataset. It is a KG-enhanced deep recommendation method that uses GAN model (Yang et al., 2018).

the probability of u_i selects v_j, and E_{v_j} is the estimation of v_j with respect to the d_T and P_s respectively. Finally, with learning the u_i - v_j optimal representations, v_j is ranked for recommendation through the P_s of G.

DKFM (Dadoun et al., 2019) translated a city-based KG via TransE to the embedding state and enhanced the representations of ideal destinations in the city to improve the performance of trip recommendations. Additionally, KTUP (Cao et al., 2019) proposed a multi-task learning approach (i.e., the unified learning of recommendation and KG completion modules). In the recommendation module,$f(\mathrm{u}_i, \mathbf{v}_j, \mathbf{p})$ is learned on u_i - v_j and v_j - v_k interlinked pairs and users' preferences to distinguish (u, v) pairs from those of (u, v^c) in the KG; where $\mathbf{p}$ shows the latent vector of u_i's preferences, and u_i - v_j and v_j - v_k pairs are extracted from $\mathscr{M}$. While in the KG completion module, $g(\boldsymbol{e}_h, \boldsymbol{r}, \boldsymbol{e}_t)$ is learned on the triplets that are extracted from KG to differentiate between the valid and invalid triplets. KTUP utilized the recommendation module to compute u_i - v_j interlinked pairs to get the preferences of users; and the KG completion module to estimate v_j - v_k potential relations in the KG. Similarly, entity2rec (Palumbo et al., 2020) proposed a property-specific KG-embedding approach to learn user-item interactions through the construction of a property-specific subgraph for items' recommendations.

4.4. Fairness aware recommendation methods

In recommendation, fairness consideration and curiosity is rapidly increasing since RS is putting outstanding impact on the selections of individuals and society due to its extraordinary contribution in decision making by streamlining the information overload (Li et al., 2021b). Thus, it is essential to optimally alleviate the issues of existing unfairness in recommendation. Although the recommendation is an unbiased ranking process concerning the fairness of the ranked options, its fairness relies on the fairness of numerous participating stakeholders. Moreover, RS is complex in tradeoff between the objectives and effectiveness to provide fair-suggestions with respect to the different applied scenarios. In this section, we categorized the unfairness in recommendation in the following three subsections.

4.4.1. Extents of fairness and biasness in recommendation

Plenty of approaches have been proposed to deal with the issues of unfairness in recommendations with respect to the extents of fairness and biasness in results. For instance, (Beutel et al., 2019) familiarized a fairness-evaluation-metric based on pairwise recommendation algorithm to assess the surety and extent of fairness in recommendation. Also, they introduced a regularization technique exploited to provoke the fairness-evaluation-metric during the training of model to well generalize the system. Similarly, (Geyik et al., 2019) introduced balancing parameters' constraints to measure the extent of biasness in the underlying interactions with respect to the potential attributes; and ranked the fairness of recommendations via a self-designed fairness aware algorithm. Similarly, (Leonhardt et al., 2018) tried to improve the fairness diversity in recommendation via measuring the unfairness identified and produced by the post-processing unit of the recommendation system. (Bonner & Vasile, 2018) Learned the proposed model on logged data that contained the recommendation results from biased strategy and predicted their own distracted recommendation results based on random users' exposures.

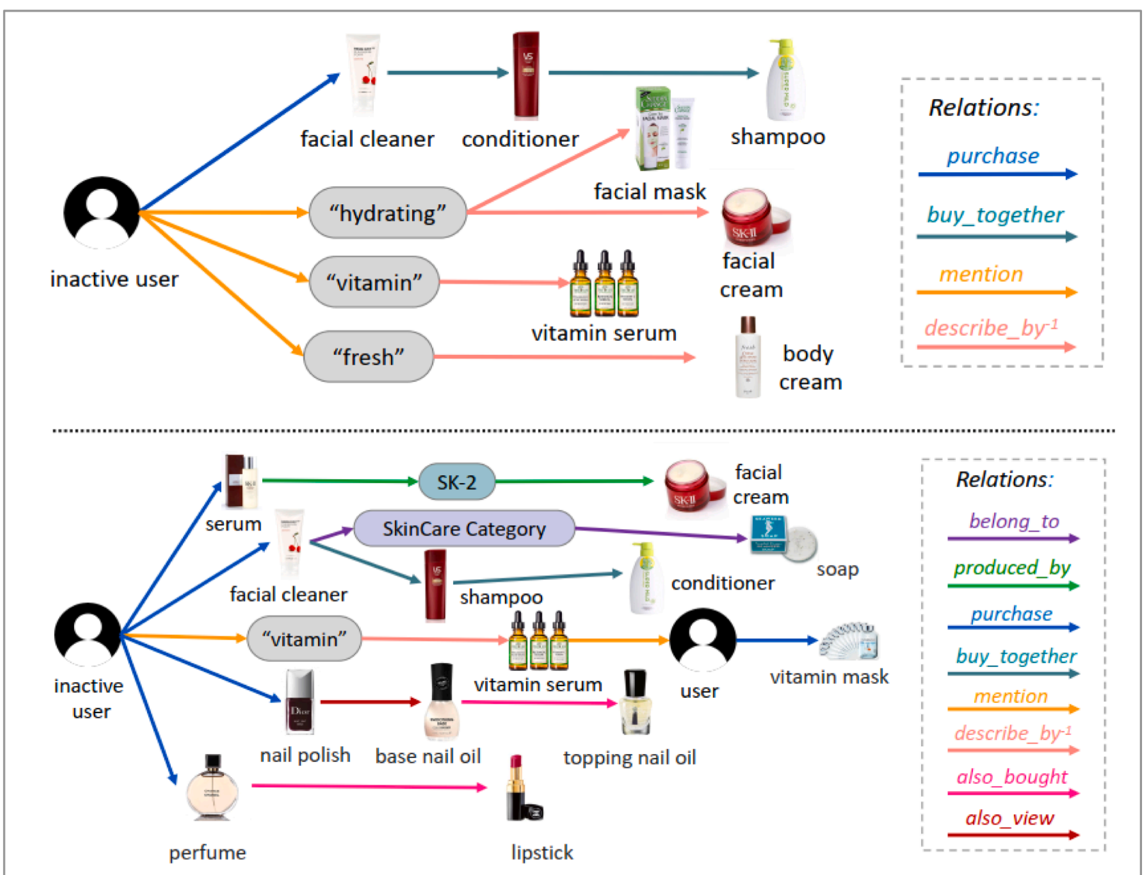

Fig. 17. The description of reasoning (represented via paths) with and without the integration of fairness-aware algorithm into the recommendation model. FAER imposed fairness constraints on users' interactions with the products' to guarantee optimally fair recommendation. For further details and explanations, please consult Sections 7.4 and 7.5 in (Fu et al., 2020).

4.4.2. Fairness against unrealistic recommendation

Researchers incorporated KG-based information into the recommendation techniques to alleviate the extents of unfairness and the biasness of unrealistic reasoning in recommendation. For instance, recently, FAER (Fu et al., 2020) discussed a scenario of unrealistic recommendations through the biased user-visibility and unfair preference-captivity; in case when there befalls a deficiency of the actual interactions between users and items. Thus, they proposed a fairness-aware ranking framework (FARF) to overcome the biases in explanation and the disparities in recommendations. An exemplary application scenario of fairness-aware recommendation is represented in Fig. 17. FAER aimed to produce the desired recommendation performance through the application of fairness-constraints with the heuristic re-ranking technique. They considered user's visibility as a user-item connection (i.e. path) represented by $\mathscr{P}$, and labeled the sequential composite relations in the user-item paths as patterns and denoted by π. The pattern π is $\{r_1{}^\circ r_2{}^\circ \cdots {}^\circ r_{|\pi|} | r_i \in R,\ i \in [|\pi|]\}$ where $^\circ$ shows the relations' composition operator. A $\mathscr{P}$ via π is $\mathscr{P}^\pi$ defined as $\mathscr{P}^\pi = \{e_0, r_1, e_1, \cdots, e_{|\pi|-1}, r_{|\pi|}, e_{|\pi|}$ such that $(e_{i-1}, r_i, e_i) \in G$ and $i \in [1, |\pi|]\}$, and π based user-item path is $\mathscr{P}^\pi_{uv}$; by satisfying $e_0 = u$ and $e_{|\pi|} = v$. Similarly, the distribution D of u-v path on u and $\mathscr{V}$ (i.e., the set of v) is $D_{u,\mathscr{V}}$; and with respect to π, it is defined as $D_{u,\mathscr{V}}(\pi) = \frac{N(\pi)}{\sum_{\pi^c \in \Pi} N(\pi^c)}$, where $N(\pi)$ is the occurrence of u-v paths w.r.t π, and $\mathscr{V}$ is the set of items written as $\left|\{\mathscr{P}^\pi_{uv} | v \in \mathscr{V}\}\right|$. To measure the degree of unfairness, FAER used Simpson's Index of Diversity (Simpson, 1949) as an evaluation metric; that calculates the probability of randomly capturing two different u-v paths (i.e., $\mathscr{P}$) belonging to the same u-v path-pattern (i.e., π). The mathematical expression to calculate the probability of getting the same path pattern in two different draws iswritten as:

$$\mathscr{T}_{uv}(\mathscr{W}, N) = 1 - \frac{\sum_{i=1}^{\mathscr{W}} n_i(n_i - 1)}{N(N-1)} \tag{19}$$

Where $\mathscr{W}$ shows the corresponding set of u to v path patterns, n_i represents the paths that belong to the i-th path pattern, and N is the total number of u-v paths initiating from u. The outcome of fairness evaluation metric (i.e., path score $\mathscr{S}_p$ and diversity score $\mathscr{S}_d$ are defined to rank the results) is used to generate the recommendations with fairness (i.e., the extent of fairness in scalar value).

4.4.3. Fair recommendation via Techniques' unification

PFGR (Xiao et al., 2020) combined the probabilistic technique with coalition game theory to guarantee the presence of fairness between the merchandising parties and the users. They treated the connections present between the groups of users and the sets of services as the relations among vertices in the graph; and imposed fairness ensuring constraints on the decision. Moreover, (Bose & Hamilton, 2019) prosed a compositional framework to impose fairness constraints on KGE to cope with the challenges of unfairness in recommendation. Typically, they claimed that the proposed approach can flexibly impose various suites of fairness constraints on the inferences. Similarly, many other approaches (Breitfuss et al., 2021; Yang & Dong, 2020) are also concerned about the indulgent assimilation of fairness into the KG-based recommendation.

4.5. Cold Start, data Sparsity, interpretability and explainability

Where the KG-based recommendation suffer from cold start and data sparsity issues, at the same time it provides the strength of interpretability and explainability with respect to the provided outcomes. In this perspectives, the concerned researchers have not only addressed the limitations of cold start and data sparsity issues through different systematic approaches, but also have highlighted the plus-points of KG-based recommendations in their respective studies. In the following subsection, we highlight few of the very concerned research examples to augment the provided explanation.

Cold start is deemed amongst the main challenges faced by the CF-based recommendation algorithms. Different researchers exploited additional supporting information to alleviate the severity of these challenges. For instance, (Jamali & Ester, 2010) generated compensation recommendations through the incorporation of social community information. Similarly, SHINE (Wang et al., 2018a) familiarized the RS with items' attribute information to handle the limitations of cold start, and (Wang et al., 2015) and (Zhang et al., 2016a) enhanced items' representations through the incorporation of multimedia and pictures into the respective learning systems. (Kang et al., 2019) introduced novel Semi Supervised Cross Domain Recommendation (SSCDR) approach to effectually model cross-domain associations in the supervised data to deal with the cold start problem. Likewise, CATN[49] (Zhao et al., 2020) learned multiple aspects of users and items from their previous interaction history and used attention mechanism to model cross-domain-based aspect correlations. DCDIR[50] (Bi et al., 2020) tackled the issue of cold start problem with respect to the arrival of new users in the data. Typically, they modelled the latent features' representations of users and items in different domains and performed a *meta*-path-based walk over the entities in the insurance product KG to locate the newly arrived users with respect to the modelled features' representations.

KGIN[51] (Wang et al., 2021) exploited item-based auxiliary knowledge to learn the intents behind the users' interactions. To enhance the interpretability and generalization capability of the model, KGIN used attention mechanism to independently model the intents as the attentive aggregation of the graph relations. (Khan et al., 2022a) introduced a conceivable solution to the issues of data sparsity in KG-based recommendation based on the data segmentation of the input datasets. In the embedding state, they segmented the input data (i.e., the interaction matrix) in six different dense parts based on the data density ratios and independently trained the model on each dense part to overcome the data sparsity problem. Similarly, (Khan et al., 2022b) proposed KGE and semantic similarity enhanced hashing-based explainable recommendation to deal with the issues of time complexity and data sparsity. (Zhang et al., 2020b) used KGE-enhanced CF-based prediction technique through DNN to predict the potential links in the KG data to describe user-item interactions. (Wang et al., 2019a) exploited the features of fine-grained services of the restaurants to classify new users based on their similar interests and behaviors with the previous users.

4.6. Discussion

"*Discovering something new based on something existent*" and "*finding user-to-item similarity through the similarity of Meta-paths*" are the incentives of path based RSs. Previously, path-based methods exploited the user-item graphs; while currently they unifies MF with the extracted Meta paths to generate recommendations based on HIN[52] (Jiang et al., 2018; Zhao et al., 2017; Shi et al., 2015; Shi et al., 2018; Yu et al., 2013a; Yu et al., 2013b; Yu et al., 2014; Luo et al., 2014). These methods enrich user-item representations through the path connectivity in KG; and need a proper domain knowledge of Meta-paths (i.e., path number, path type, length, etc.) for this purpose – a drawback of path-based methods. To handle this issue, (Ma et al., 2019) proposed RuleRec to automatically exploit the external KGs to provide recommendation-rules. Recently, different DL-based approaches (Song et al., 2019a; Huang et al., 2019a; Xian et al., 2019; Wang et al., 2019g; Sun et al., 2018b; Hu et al., 2018; Zhang et al., 2016a) proposed path-embedding (and automatic extraction of the prominent paths from KG) to inter-link the indistinguishable (i.e., potentially similar, identical) users and items for

49 Cross-domain recommendation framework via Aspect Transfer Network.
50 Deep Cross Domain Insurance Recommendation System.
51 Knowledge Graph-based Intent Network.
52 Heterogeneous Information Network.

recommendations. Moreover, (Fu et al., 2020) proposed FARF to provide fairness-aware recommendations.

Exploitation of KG-based side information varied with the variation in application requirements. For instance, (Zhang et al., 2018c; Dadoun et al., 2019; Wang et al., 2018a; Palumbo et al., 2017; Ai et al., 2018) directly modeled user preferences through the inclusion of user information to KG and the creation of user-item sub-graphs. Similarly, node2vec (Palumbo et al., 2018) proposed feedback-based KG-embedding approach; and entity2rec (Palumbo et al., 2020) developed a property-specific KG-embedding approach to recommend items. (Zhang et al., 2016a; Yang et al., 2018; Cao et al., 2019; Joseph & Jiang, 2019; Xin et al., 2019; Ye et al., 2019; Wang et al., 2018b; Wang et al., 2019d; Huang et al., 2018) proposed different KGE approaches to enhance items' representations through the external KReps and construct local subgraphs for implementation. Items' representations are used to formulate users' representations more accurately.

(Yang et al., 2018; Ye et al., 2019) improved the embedding of entities to improve the performance of recommendations. Moreover, (Piao & Breslin, 2018; Xin et al., 2019; Cao et al., 2019; Wang et al., 2019d) proposed multi-task learning approaches, i.e., the unified learning of recommendation and KG completion modules. Recently, (Yang et al., 2020) exploited KG-based local and non-local context information of entities for recommendations, and (Wang et al., 2020b) utilized RL to discover high quality negatives in KG to control negative sampling in recommendations.

5. Future Research Directions

"Room is always there for improvement!" Although researchers did their best to enhance the performance of KG-based recommendation methods, considerable challenges also pop-up with the advancement in internet technology and online data with the passage of time. Therefore, by taking inspiration from (Guo et al., 2020), we are eager to identify current challenges and describe potential research directions for the future researchers.

5.1. KGE enhancement via Interaction's collection and noise filtration

Explicit user feedback is one of the major sources of acquiring user preferences for the similar potential items. If systems provide multiple options to the users to give feedback, the process of data collection can be made convenient and sophisticated. Similarly, crowdsourcing is an efficient technique of data collection for future data processing and decision making. Although it is normally utilized by the classical recommender systems, it can be effectively incorporated as a central component of information collection and filtration into the KG-based recommendation methods. To enhance the recommendation performance, entity-relevance (i.e., semantic similarity based information filtration to avoid possible emergence of noise in the underlying information during the construction of subgraph) is also a future research direction of KG-based recommendation methods.

5.2. KGE enhancement via language modeling

The integration of external knowledge into the models of language representation enrich the local representations of text with respect to the knowledge. For instance, DKN (Wang et al., 2018b) concatenated the embeddings of entity and text to acquire the final representations of News. STCKA (Chen et al., 2019) classified short-text via enhancing its semantic representations with the knowledge extracted from YAGO, and ERNIE (Zhang et al., 2019) – a relation classifier – aggregated the language information with the knowledge extracted from Wikidata to enrich the language representations. Therefore, it is understood that the enriched representation learning enhances the performance of recommendation. The enriched language representation is a future research dimension of KG-based data processing through the NLP operations. Similarly, translation distance models and semantic matching approaches belong to KGE under the applied constraints. Although a lot of working environments such as data sourcing, information targeting, similarity matching, model structuring, graph building, etc., are available, there is lack of a clear and comprehensive scenario that deal with the available environments via the KGE approach. Under the applied constraints, suggesting (i.e., designing and implementing) comprehensive working environments to deal with the above mentioned application-scenarios with respect to the data provision (i.e., database) through the KGE approach is a future research perspective.

5.3. KGE enhancement via Multi-Domain interactional data

Either the interacting platforms cannot equally contribute to the data sharing across domains or all of the domains do not possess equal amounts of data. For example, Books or products ratings on amazon are greater than any other platform. Thus, data-vise rich source-domains can share the interactional data with the comparatively poor target-domains via transfer-learning techniques. KerKT (Zhang et al., 2018b) proposed matrix factorization method to find user-item similarities across different domains with the kernel induced knowledge transfer technique. Similarly, PPGN (Zhao et al., 2019a) used their user-item interaction KG, constructed of users and items from dissimilar domains, in a cross domain system. Though PPGN dominantly outperformed the state of the art models in case of user-item interaction graph, it remained unable to disclose or deal with the other connections of user or item beside their interaction KG. Therefore, handling user-item un-interacted relations in the case of PPGN model; and integrating user-item information from different domains into the user-item interaction KG to enhance the performance of cross-domain-based RSs; are the future research dimensions.

5.4. KGE enhancement via continual learning

Although GNN and GCN performed terrifically well, these are the static preference acquisition approaches. Currently, recommendation about the Fast Influencing Interests (FII) like friends and family greetings, reviews/notifications about the items of concern, Twitter hashtags, social media, social gamming, and news scenarios need real-time recommendation and response. Dynamic Graph Network is a solution to dynamically incorporate current information into the underlying data; and Continual Graph Learning and Continuous KGE enhancement are the examples of this scenario. (Song et al., 2019b) acquired rapidly varying interests of the users from their past interests as well as from long and short term interests of their friends through DGAN[53] model. Therefore, the removal of irrelevant information from the received data, construction and continual enhancement of specialized KGs in KGE mode, designing of real-time retrieval & responsive techniques, and their aggregation for the acquisition of dynamic feedback are the future research dimensions.

5.5. Incorporation of Users' side information into the KGE

In the current research trends of KG-based recommendation, most of the researchers used items' side information to create their knowledge graph for recommendations; but there exist few approaches that incorporated users' side information as well into the KGE. It implies that users' side information (e.g., user's demography, profile information, social network, friends circle, types of interest, past interaction data, etc.) can also be included to the KGs and KGE. For instance, (Sha et al., 2019) highlighted user connections with other entities in the KG and proved that the integration of user side information is an effective mechanism to enhance the performance of recommendation. Similarly,

[53] Dynamic Graph Attention Network.

(Fan et al., 2019) outperformed the traditional CF-based system with user side information by separately representing user-user and user-item relations in the social and interaction KGs, respectively; through GNN. Moreover, (Sang et al., 2021) outperformed the traditional CF-based systems with neural CF-based systems. Therefore, the integration of users' side information into the KG; and the upgradation of traditional filtering-based systems to work on KGs can be considered as the future research dimensions.

5.6. Triplet-based prediction of the unknown entity or relation via the known information

KG-based recommendation and KG-based link-prediction (i.e., among entities in KG) methods are mutually analogous with respect to the uni-task learning, but there are some reasons that drag them to the multi-task learning approaches as well. Formally, it is evident that losing the information is losing the performance. If there exist such triplets that are discarded (or missed) by the model during the construction of sub-graph, the process will lead to exclude (loose) the information they covered. Therefore, (Wang et al., 2019d) and (Xin et al., 2019) jointly learned the recommendation module with KGE and item-relations regularization, respectively. Similarly, (Li et al., 2019c; Cao et al., 2019) jointly learned the modules of KG-completion and recommendation and demonstrated that their approaches are effective. Therefore, tracking the link leading from a known head-entity to an unknown tail-entity can be utilized with the path-information to guess the destination (i.e., the missing tail-entity). Similarly, the unknown head-entity or the relation can also be predicted based on the known-information of other two (i.e., the tail entity and the relation, or the head and the tail entities). Moreover, transferring the information from the entity-classification, the resolution and other KG-based operations to improve the recommendations is a future research dimension. Furthermore, Path Tracking Systems (Dogan et al., 2020) are somehow similar to the path-crawling techniques (i.e., path-based methods in KG-based RSs) with respect to the working mechanism and implementation. Consequently, we hope and dare to narrate that the KGs can be incorporated into the Path Tracking Systems in future.

6. Conclusion

We explored the domain of KG-based recommendation and presented it in a well-categorized way to express how knowledge graph offers interpretability and provide side information to RSs to enhance the performance. In this paper, we defined KG, RSs, KG-based RSs, presented an overview of KReps, and categorically discussed the benchmark datasets and recommendation methods with respect to the application scenarios. We also focused on the construction of local subgraphs, either directly from the independent datasets[54] or using the datasets with the information extracted from the external KReps to enrich the entity representations. Finally, we recommended some future research dimensions to assist the fresh researchers in this domain. Subsequent to devouring the literature, we conclusively narrate that KG-based recommendation not only generate accurate, personalized and explainable recommendations, but also provide potential and productive information to the end users. We hope this paper can help the fresh researchers to better understand the domain.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

[54] These datasets are used by the proposed models during the experiments.

Acknowledgements

The work was supported in part by the Basic Research Program of Jiangsu Province (BK20191274) and the National Natural Science Foundation of China (62176121 and 61772269).

References

Ai, Q., Azizi, V., Chen, X., & Zhang, Y. (2018). Learning heterogeneous knowledge base embeddings for explainable recommendation. *Algorithms, 11*, 137.
Alam, M., Recupero, D. R., Mongiovi, M., Gangemi, A., & Ristoski, P. (2017). Reconciling event-based knowledge through rdf2vec. In *HybridSemStats@ ISWC*.
Alonso-Maturana, R., Alvarado-Cortes, E., Ĺopez-Sola, S., Martınez-Losa, M. O., & Hermoso-González, P. (2018). La rioja turismo: The construction and exploitation of a queryable tourism knowledge graph. In *International Conference on Web Engineering* (pp. 213–220). Springer.
Asghar, N. (2016). Yelp dataset challenge: Review rating prediction. *arXiv preprint arXiv:1605.05362*.
Basile, P., Musto, C., de Gemmis, M., Lops, P., Narducci, F., & Semeraro, G. (2014). Aggregation strategies for linked open data-enabled recommender systems. *11th ESWC*.
Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C. (2019). A review on deep learning for recommender systems: Challenges and remedies. *Artificial Intelligence Review, 52*, 1–37.
Belleau, F., Nolin, M.-A., Tourigny, N., Rigault, P., & Morissette, J. (2008). Bio2rdf: Towards a mashup to build bioinformatics knowledge systems. *Journal of Biomedical Informatics, 41*, 706–716.
Bellini, V., Schiavone, A., Di Noia, T., Ragone, A., & Di Sciascio, E. (2018). Computing recommendations via a knowledge graph-aware autoencoder.
Beutel, A., Chen, J., Doshi, T., Qian, H., Wei, L., Wu, Y., et al. (2019). Fairness in recommendation ranking through pairwise comparisons. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2212–2220).
Bi, Y., Song, L., Yao, M., Wu, Z., Wang, J., & Xiao, J. (2020). Dcdir: A deep cross-domain recommendation system for cold start users in insurance domain. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval* (pp. 1661–1664).
Bizer, C., Heath, T., & Berners-Lee, T. (2011). Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts* (pp. 205–227). IGI global.
Blanco, R., Cambazoglu, B. B., Mika, P., & Torzec, N. (2013). In *Entity recommendations in web search* (pp. 33–48). Springer.
Bollacker, K., Evans, C., Paritosh, P., Sturge, T., & Taylor, J. (2008). Freebase: A collaboratively created graph database for structuring human knowledge. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 1247–1250).
Bonner, S., & Vasile, F. (2018). Causal embeddings for recommendation. In *Proceedings of the 12th ACM conference on recommender systems* (pp. 104–112).
Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In *Advances in neural information processing systems* (p. 26).
Bose, A., & Hamilton, W. (2019). Compositional fairness constraints for graph embeddings. In *International Conference on Machine Learning* (pp. 715–724). PMLR.
Brämer, G. R. (1988). International statistical classification of diseases and related health problems. tenth revision. *World health statistics quarterly. Rapport trimestriel de statistiques sanitaires mondiales, 41*, 32–36.
Brauner, D. F., Intrator, C., Freitas, J. C., & Casanova, M. A. (2007). An instancebased approach for matching export schemas of geographical database web services. *In GeoInfo*, 109–120.
Breitfuss, A., Errou, K., Kurteva, A., & Fensel, A. (2021). Representing emotions with knowledge graphs for movie recommendations. *Future Generation Computer Systems*.
Burke, R. (2022). Personalized recommendation of pois to people with autism: technical perspective. *Communications of the ACM, 65*, 100–100.
Cai, H., Zheng, V. W., & Chang, K.-C.-C. (2018). A comprehensive survey of graph embedding: Problems, techniques, and applications. *IEEE Transactions on Knowledge and Data Engineering, 30*, 1616–1637.
Cano, E., & Morisio, M. (2017). Hybrid recommender systems: A systematic literature review. *Intelligent Data Analysis, 21*, 1487–1524.
Cao, E., Wang, D., Huang, J., & Hu, W. (2020). Open knowledge enrichment for long-tail entities. In *Proceedings of The Web Conference 2020* (pp. 384–394).
Cao, Y., Wang, X., He, X., Hu, Z., & Chua, T.-S. (2019). Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences. In *The world wide web conference* (pp. 151–161).
Cardoso, S. D., Da Silveira, M., & Pruski, C. (2020). Construction and exploitation of an historical knowledge graph to deal with the evolution of ontologies. *Knowledge-Based Systems, 194*, Article 105508.
Carlson, A., Betteridge, J., Wang, R. C., Hruschka, E. R., Jr, & Mitchell, T. M. (2010). Coupled semi-supervised learning for information extraction. In *Proceedings of the third ACM international conference on Web search and data mining* (pp. 101–110).
Catherine, R., & Cohen, W. (2016). Personalized recommendations using knowledge graphs: A probabilistic logic programming approach. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 325–332).
Chah, N. (2017). Freebase-triples: A methodology for processing the freebase data dumps. *arXiv preprint arXiv:1712.08707*.

Chen, C., Zhang, M., Ma, W., Liu, Y., & Ma, S. (2020a). Jointly non-sampling learning for knowledge graph enhanced recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 189–198).

Chen, J., Hu, Y., Liu, J., Xiao, Y., & Jiang, H. (2019). Deep short text classification with knowledge powered attention. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 6252–6259). volume 33.

Chen, R., Hua, Q., Chang, Y.-S., Wang, B., Zhang, L., & Kong, X. (2018). A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks. *IEEE Access, 6*, 64301–64320.

Chen, X., Jia, S., & Xiang, Y. (2020b). A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications, 141*, Article 112948.

Chen, X., Qin, Z., Zhang, Y., & Xu, T. (2016). Learning to rank features for recommendation over multiple categories. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 305–314).

Dadoun, A., Troncy, R., Ratier, O., & Petitti, R. (2019). Location embeddings for next trip recommendation. In *Companion Proceedings of The 2019 World Wide Web Conference* (pp. 896–903).

Das, R., Neelakantan, A., Belanger, D., & McCallum, A. (2016). Chains of reasoning over entities, relations, and text using recurrent neural networks. *arXiv preprint arXiv: 1607.01426*.

De Gemmis, M., Lops, P., Musto, C., Narducci, F., & Semeraro, G. (2015). Semanticsaware content-based recommender systems. In *Recommender systems handbook* (pp. 119–159). Springer.

De Sa, C., Ratner, A., Ŕe, C., Shin, J., Wang, F., Wu, S., & Zhang, C. (2016). Deepdive: Declarative knowledge base construction. *ACM SIGMOD Record, 45*, 60–67.

Dehmer, M., & Varmuza, K. (2015). A comparative analysis of the tanimoto index and graph edit distance for measuring the topological similarity of trees. *Applied Mathematics and Computation, 259*, 242–250.

Deldjoo, Y., Constantin, M. G., Ionescu, B., Schedl, M., & Cremonesi, P. (2018). Mmtf-14k: A multifaceted movie trailer feature dataset for recommendation and retrieval. In *Proceedings of the 9th ACM Multimedia Systems Conference* (pp. 450–455).

Di Noia, T., Cantador, I., & Ostuni, V. C. (2014). Linked open data-enabled recommender systems: Eswc 2014 challenge on book recommendation. In *Semantic Web Evaluation Challenge* (pp. 129–143). Springer.

Ding, C. H., Li, T., & Jordan, M. I. (2008). Convex and semi-nonnegative matrix factorizations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*, 45–55.

Ding, Y., Ma, Y., Wong, W. K., & Chua, T.-S. (2021). Leveraging two types of global graph for sequential fashion recommendation. In *Proceedings of the 2021 International Conference on Multimedia Retrieval* (pp. 73–81).

Dogan, M. U., Guvenc, U., & Elmas, C. (2020). Genetic pi based model and path tracking control of four traction electrical vehicle. *Electrical Engineering, 102*, 2059–2073.

Dong, X., Gabrilovich, E., Heitz, G., Horn, W., Lao, N., Murphy, K., et al. (2014). Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 601–610).

Dong, Y., Chawla, N. V., & Swami, A. (2017). metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 135–144).

Dooms, S., Belloǵın, A., Pessemier, T. D., & Martens, L. (2016). A framework for dataset benchmarking and its application to a new movie rating dataset. *ACM Transactions on Intelligent Systems and Technology (TIST), 7*, 1–28.

Dragisic, Z., Eckert, K., Euzenat, J., Faria, D., Ferrara, A., Granada, R. L., Ivanova, V., Jiḿenez-Ruiz, E., Kempf, A., Lambrix, P. et al. (2014). Results of the ontology alignment evaluation initiative 2014. In *11th International Workshop on Ontology Matching co-located with the 13th ISWC (OM 2014)* (pp. 61–104). volume 1317.

Ehrlinger, L., & Wöß, W. (2016). Towards a definition of knowledge graphs. semantics (posters, demos, success). *12th International Conference on Semantic Systems*.

Ernst, P., Meng, C., Siu, A., & Weikum, G. (2014). Knowlife: a knowledge graph for health and life sciences. In *2014 IEEE 30th International Conference on Data Engineering* (pp. 1254–1257). IEEE.

Fabian, M., Gjergji, K., Gerhard, W., et al. (2007). Yago: A core of semantic knowledge unifying wordnet and wikipedia. In *16th International World Wide Web Conference* (pp. 697–706).

Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., & Yin, D. (2019). Graph neural networks for social recommendation. In *The World Wide Web Conference* (pp. 417–426).

Fang, Y., Lin, W., Zheng, V. W., Wu, M., Chang, K.-C.-C., & Li, X.-L. (2016). In *Semantic proximity search on graphs with metagraph-based learning* (pp. 277–288). IEEE.

Färber, M. (2019). The microsoft academic knowledge graph: a linked data source with 8 billion triples of scholarly data. In *International Semantic Web Conference* (pp. 113–129). Springer.

Färber, M., & Rettinger, A. (2018). Which knowledge graph is best for me? *arXiv preprint arXiv:1809.11099*.

Fu, T.-y., Lee, W.-C., & Lei, Z. (2017). Hin2vec: Explore meta-paths in heterogeneous information networks for representation learning. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 1797–1806).

Fu, Z., Xian, Y., Gao, R., Zhao, J., Huang, Q., Ge, Y., et al. (2020). Fairness-aware explainable recommendation over knowledge graphs. In *In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 69–78).

Gardner, M., Talukdar, P., Krishnamurthy, J., & Mitchell, T. (2014). Incorporating vector space similarity in random walk inference over knowledge bases. In *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 397–406).

Geyik, S. C., Ambler, S., & Kenthapadi, K. (2019). Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. In *In Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining* (pp. 2221–2231).

Giunchiglia, F., Maltese, V., Farazi, F., & Dutta, B. (2010). Geowordnet: a resource for geo-spatial applications. In *Extended Semantic Web Conference* (pp. 121– 136). Springer.

Gomez-Perez, J. M., Pan, J. Z., Vetere, G., & Wu, H. (2017). Enterprise knowledge graph: An introduction. In *Exploiting linked data and knowledge graphs in large organisations* (pp. 1–14). Springer.

Goyal, N., Sachdeva, N., Choudhary, V., Kar, R., Kumaraguru, P., & Rajput, N. (2019). Con2kg-a large-scale domain-specific knowledge graph. In *In Proceedings of the 30th ACM Conference on Hypertext and Social Media* (pp. 287–288).

Grover, A., & Leskovec, J. (2016). node2vec: Scalable feature learning for networks. In *In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 855–864).

Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., et al. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*.

Gutlein, M., Frank, E., Hall, M., & Karwath, A. (2009). Large-scale attribute selection using wrappers. In *2009 IEEE symposium on computational intelligence and data mining* (pp. 332–339). IEEE.

Hakkani-Tur, D., Celikyilmaz, A., Heck, L., Tur, G., & Zweig, G. (2014). Probabilistic enrichment of knowledge graph entities for relation detection in conversational understanding. *In Proceedings of INTERSPEECH*.

Han, X., Yu, P., Liu, Z., Sun, M., & Li, P. (2018). Hierarchical relation extraction with coarse-to-fine grained attention. In *In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 2236–2245).

Harper, F. M., & Konstan, J. A. (2015). The movielens datasets: History and context. *Acm Transactions on Interactive Intelligent Systems (tiis), 5*, 1–19.

Hoffart, J., Suchanek, F. M., Berberich, K., Lewis-Kelham, E., De Melo, G., & Weikum, G. (2011). Yago2: Exploring and querying world knowledge in time, space, context, and many languages. In *In Proceedings of the 20th international conference companion on World wide web* (pp. 229–232).

Hoyt, C. T., Domingo-Fernández, D., Aldisi, R., Xu, L., Kolpeja, K., Spalek, S., Wollert, E., Bachman, J., Gyori, B. M., Greene, P. et al. (2019). Re-curation and rational enrichment of knowledge graphs in biological expression language. *Database, 2019*.

Hsieh, M.-Y., Chou, W.-K., & Li, K.-C. (2017). Building a mobile movie recommendation service by user rating and app usage with linked data on hadoop. *Multimedia Tools and Applications, 76*, 3383–3401.

Hu, B., Shi, C., Zhao, W. X., & Yu, P. S. (2018). Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In *In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1531–1540).

Hu, Z., Huang, P., Deng, Y., Gao, Y., & Xing, E. (2015). Entity hierarchy embedding. In *In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 1292–1300).

Huang, J., Zhao, W. X., Dou, H., Wen, J.-R., & Chang, E. Y. (2018). Improving sequential recommendation with knowledge-enhanced memory networks. In *In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 505–514).

Huang, S., & Wan, X. (2013). Akminer: Domain-specific knowledge graph mining from academic literatures. In *International Conference on Web Information Systems Engineering* (pp. 241–255). Springer.

Huang, X., Fang, Q., Qian, S., Sang, J., Li, Y., & Xu, C. (2019a). Explainable interaction-driven user modeling over knowledge graph for sequential recommendation. In *In Proceedings of the 27th ACM International Conference on Multimedia* (pp. 548–556).

Huang, X., Zhang, J., Li, D., & Li, P. (2019b). Knowledge graph embedding based question answering. In *In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 105–113).

Huang, Y., Li, M., & Wu, Y. (2019c). Kkbox's music recommendation.

Ivarsson, J., & Lindgren, M. (2016). Movie recommendations using matrix factorization.

Jagvaral, B., Lee, W.-K., Roh, J.-S., Kim, M.-S., & Park, Y.-T. (2020). Path-based reasoning approach for knowledge graph completion using cnn-bilstm with attention mechanism. *Expert Systems with Applications, 142*, Article 112960.

Jamali, M., & Ester, M. (2010). A matrix factorization technique with trust propagation for recommendation in social networks. In *In Proceedings of the fourth ACM conference on Recommender systems* (pp. 135–142).

Ji, G., He, S., Xu, L., Liu, K., & Zhao, J. (2015). Knowledge graph embedding via dynamic mapping matrix. In *In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 687–696).

Ji, G., Liu, K., He, S., & Zhao, J. (2017). Distant supervision for relation extraction with sentence-level attention and entity descriptions. In *Proceedings of the AAAI Conference on Artificial Intelligence*. volume 31.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*.

Jia, Y., Wang, Y., Lin, H., Jin, X., & Cheng, X. (2016). Locally adaptive translation for knowledge graph embedding. *In Thirtieth AAAI conference on artificial*. intelligence.

Jiang, H., Cui, L., Xu, Z., Yang, D., Chen, J., Li, C., Liu, J., Liang, J., Wang, C., Xiao, Y. et al. (2019). Relation extraction using supervision from topic knowledge of relation labels. In *IJCAI* (pp. 5024–5030).

Jiang, Z., Liu, H., Fu, B., Wu, Z., & Zhang, T. (2018). Recommendation in heterogeneous information networks based on generalized random walk model and bayesian personalized ranking. In *In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 288–296).

Joseph, K., & Jiang, H. (2019). Content based news recommendation via shortest entity distance over knowledge graphs. In *In Companion Proceedings of The 2019 World Wide Web Conference* (pp. 690–699).

Kang, S., Hwang, J., Lee, D., & Yu, H. (2019). Semi-supervised learning for crossdomain recommendation to cold-start users. In *In Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 1563–1572).

Kanyundo, A., & Kapondera, S. K. (2016). Relieved at last: Cataloguing with librarything. *In 22nd Standing Conference of the Eastern, Central and Southern Africa Librarians and Institutions (SCECSAL)*.

Khan, N., Ma, Z., Ullah, A., & Polat, K. (2022a). Dca-iomt: Knowledge graph embedding-enhanced deep collaborative alerts-recommendation against covid19. *IEEE Transactions on Industrial Informatics*. https://doi.org/10.1109/TII.2022. 3159710

Khan, N., Ma, Z., Yan, L., & Ullah, A. (2022b). Hashing-based semantic relevance attributed knowledge graph embedding enhancement for deep probabilistic recommendation. *Applied Intelligence*, 1–26.

A. Kilgarriff Wordnet: An electronic lexical database 2000.

Kontokostas, D., Bratsas, C., Auer, S., Hellmann, S., Antoniou, I., & Metakides, G. (2012). Internationalization of linked data: The case of the greek dbpedia edition. *Journal of Web Semantics*, *15*, 51–61.

Kouki, P., Schaffer, J., Pujara, J., O'Donovan, J., & Getoor, L. (2019). Personalized explanations for hybrid recommender systems. In *In Proceedings of the 24th International Conference on Intelligent User Interfaces* (pp. 379–390).

Krompaß, D., Baier, S., & Tresp, V. (2015). Type-constrained representation learning in knowledge graphs. In *International semantic web conference* (pp. 640–655). Springer.

Lao, N., & Cohen, W. W. (2010). Relational retrieval using a combination of pathconstrained random walks. *Machine Learning, 81*, 53–67.

Leacock, C., & Chodorow, M. (1998). Combining local context and wordnet similarity for word sense identification. *WordNet: An Electronic Lexical database, 49*, 265–283.

Lee, D., & Hosanagar, K. (2014). Impact of recommender systems on sales volume and diversity.

Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., et al. (2015). Dbpedia–a largescale, multilingual knowledge base extracted from wikipedia. *Semantic web, 6*, 167–195.

Lenat, D. B. (1995). Cyc: A large-scale investment in knowledge infrastructure. *Communications of the ACM, 38*, 33–38.

Leonhardt, J., Anand, A., & Khosla, M. (2018). User fairness in recommender systems. In *In Companion Proceedings of the The Web Conference 2018* (pp. 101–102).

Leung, J. K., Griva, I., & Kennedy, W. G. (2020a). Making use of affective features from media content metadata for better movie recommendation making. arXiv preprint arXiv:2007.00636.

Leung, J. K., Griva, I., & Kennedy, W. G. (2020b). Text-based emotion aware recommender. arXiv preprint arXiv:2007.01455.

Li, H., Liu, Y., Mamoulis, N., & Rosenblum, D. S. (2019a). Translation-based sequential recommendation for complex users on sparse data. *IEEE Transactions on Knowledge and Data Engineering, 32*, 1639–1651.

Li, L., Wang, P., Wang, Y., Jiang, J., Tang, B., Yan, J., Wang, S., & Liu, Y. (2019b). Prtransh: embedding probabilistic medical knowledge from real world emr data.

Li, L., Wang, P., Yan, J., Wang, Y., Li, S., Jiang, J., et al. (2020). Real-world data medical knowledge graph: Construction and applications. *Artificial Intelligence in Medicine, 103*, Article 101817.

Li, Q., Tang, X., Wang, T., Yang, H., & Song, H. (2019c). Unifying task-oriented knowledge graph learning and recommendation. *IEEE Access, 7*, 115816–115828.

Li, X., Lyu, M., Wang, Z., Chen, C.-H., & Zheng, P. (2021a). Exploiting knowledge graphs in industrial products and services: A survey of key aspects, challenges, and future perspectives. *Computers in Industry, 129*, Article 103449.

Li, Y., Ge, Y., & Zhang, Y. (2021b). Tutorial on fairness of machine learning in recommender systems. *SIGIR*.

Lin, H., Liu, Y., Wang, W., Yue, Y., & Lin, Z. (2017). Learning entity and relation embeddings for knowledge resolution. *Procedia Computer Science, 108*, 345–354.

Lin, Y., Liu, Z., & Sun, M. (2016). Knowledge representation learning with entities, attributes and relations. *Ethnicity, 1*, 41–52.

Lin, Y., Liu, Z., Sun, M., Liu, Y., & Zhu, X. (2015). Learning entity and relation embeddings for knowledge graph completion. In Twenty-ninth AAAI conference on artificial intelligence.

Liu, D., Bai, T., Lian, J., Zhao, X., Sun, G., Wen, J.-R., & Xie, X. (2019). News graph: An enhanced knowledge graph for news recommendation. In KaRS@ CIKM (pp. 1–7).

Luo, C., Pang, W., Wang, Z., & Lin, C. (2014). Hete-cf: Social-based collaborative filtering recommendation using heterogeneous relations. In 2014 IEEE International Conference on Data Mining (pp. 917–922). IEEE.

Ma, W., Zhang, M., Cao, Y., Jin, W., Wang, C., Liu, Y., et al. (2019). Jointly learning explainable rules for recommendation with knowledge graph. In *In The World Wide Web Conference* (pp. 1210–1221).

Ma, Y., Cambria, E., & Gao, S. (2016). Label embedding for zero-shot fine-grained named entity typing. In *In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 171–180).

Mahdisoltani, F., Biega, J., & Suchanek, F. (2014). Yago3: A knowledge base from multilingual wikipedias. *In 7th biennial conference on innovative data systems research. CIDR Conference*.

Mauro, N., Ardissono, L., & Cena, F. (2020). Personalized recommendation of pois to people with autism. In Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (pp. 163–172).

McAuley, J., Targett, C., Shi, Q., & Van Den Hengel, A. (2015). Image-based recommendations on styles and substitutes. In *In Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval* (pp. 43–52).

Meusel, R., Petrovski, P., & Bizer, C. (2014). The webdatacommons microdata, rdfa and microformat dataset series. In International Semantic Web Conference (pp. 277–292). Springer.

Michel, F., Gandon, F., Ah-Kane, V., Bobasheva, A., Cabrio, E., Corby, O., Gazzotti, R., Giboin, A., Marro, S., Mayer, T. et al. (2020). Covid-on-the-web: Knowledge graph and services to advance covid-19 research. In International Semantic Web Conference (pp. 294–310). Springer.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.

Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., et al. (2018). Never-ending learning. *Communications of the ACM, 61*, 103–115.

Moler, C., & Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Review, 45*, 3–49.

Monnin, P., Raïssi, C., Napoli, A., & Coulet, A. (2019). Knowledge reconciliation with graph convolutional networks: Preliminary results. In DL4KG2019-Workshop on Deep Learning for Knowledge Graphs 2377.

Morin, F., & Bengio, Y. (2005). Hierarchical probabilistic neural network language model. In International workshop on artificial intelligence and statistics (pp. 246–252). PMLR.

Morsey, M., Lehmann, J., Auer, S., Stadler, C., & Hellmann, S. (2012). Dbpedia and the live extraction of structured data from wikipedia. *Program*.

Nakashole, N., Theobald, M., & Weikum, G. (2011). Scalable knowledge harvesting with high precision and high recall. In *In Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 227–236).

Neelakantan, A., Roth, B., & McCallum, A. (2015). Compositional vector space models for knowledge base completion. arXiv preprint arXiv:1504.06662.

Nguyen, Q. N., Duong-Trung, N., Le Ha, D. N., Ha, X. S., Phan, T. T., Pham, H. X., et al. (2020). Movie recommender systems made through tag interpolation. In *In Proceedings of the 4th international conference on machine learning and soft computing* (pp. 154–158).

Nilsson, N. J. (1996). Artificial intelligence: A modern approach. *Artificial intelligence, 82*.

Niu, F., Zhang, C., Ȓe, C., & Shavlik, J. W. (2012). Deepdive: Web-scale knowledgebase construction using statistical learning and inference. VLDS, 12, 25–28.

Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A., & Taylor, J. (2019). Industry-scale knowledge graphs: Lessons and challenges: Five diverse technology companies show how it's done. *Queue, 17*, 48–75.

Nurdiati, S., & Hoede, C. (2008). 25 years development of knowledge graph theory: The results and the challenge. *Memorandum, 1876*, 1–10.

Palumbo, E., Monti, D., Rizzo, G., Troncy, R., & Baralis, E. (2020). entity2rec: Property-specific knowledge graph embeddings for item recommendation. *Expert Systems with Applications, 151*, Article 113235.

Palumbo, E., Rizzo, G., & Troncy, R. (2017). Entity2rec: Learning user-item relatedness from knowledge graphs for top-n item recommendation. In *In Proceedings of the eleventh ACM conference on recommender systems* (pp. 32–36).

Palumbo, E., Rizzo, G., Troncy, R., Baralis, E., Osella, M., & Ferro, E. (2018). Knowledge graph embeddings with node2vec for item recommendation. In European Semantic Web Conference (pp. 117–120). Springer.

Park, J., Cho, Y., Lee, H., Choo, J., & Choi, E. (2021). Knowledge graph-based question answering with electronic health records. In Machine Learning for Healthcare Conference (pp. 36–53). PMLR.

Paulheim, H. (2017). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic web, 8*, 489–508.

Pellissier Tanon, T., Vrandečić, D., Schaffert, S., Steiner, T., & Pintscher, L. (2016). From freebase to wikidata: The great migration. In *In Proceedings of the 25th international conference on world wide web* (pp. 1419–1428).

Penha, G., & Hauff, C. (2020). What does bert know about books, movies and music? probing bert for conversational recommendation. In *In Fourteenth ACM Conference on Recommender Systems* (pp. 388–397).

Piao, G., & Breslin, J. G. (2018). Transfer learning for item recommendations and knowledge graph completion in item related domains via a co-factorization model. In European Semantic Web Conference (pp. 496–511). Springer.

Ponzetto, S. P., & Strube, M. (2007). Knowledge derived from wikipedia for computing semantic relatedness. *Journal of Artificial Intelligence Research, 30*, 181–212.

Qian, R. (2013). Understand your world with bing. Bing search blog, Mar.

Qin, C., Zhu, H., Zhu, C., Xu, T., Zhuang, F., Ma, C., et al. (2019). Duerquiz: A personalized question recommender system for intelligent job interview. In *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2165–2173).

Qin, T., & Liu, T.-Y. (2013). Introducing letor 4.0 datasets. arXiv preprint arXiv: 1306.2597.

Qu, Y., Bai, T., Zhang, W., Nie, J., & Tang, J. (2019). An end-to-end neighborhoodbased interaction model for knowledge-enhanced recommendation. In *In Proceedings of the 1st International Workshop on Deep Learning Practice for HighDimensional Sparse Data* (pp. 1–9).

Rafailidis, D. (2019). Bayesian deep learning with trust and distrust in recommendation systems. In 2019 IEEE/WIC/ACM International Conference on Web Intelligence (WI) (pp. 18–25). IEEE.

Sang, L., Xu, M., Qian, S., & Wu, X. (2021). Knowledge graph enhanced neural collaborative recommendation. *Expert Systems with Applications, 164*, Article 113992.

Schedl, M. (2016). The lfm-1b dataset for music retrieval and recommendation. In *In Proceedings of the 2016 ACM on international conference on multimedia retrieval* (pp. 103–110).

Sha, X., Sun, Z., & Zhang, J. (2019). Attentive knowledge graph embedding for personalized recommendation. arXiv preprint arXiv:1910.08288.
Shao, B., Li, X., & Bian, G. (2021). A survey of research hotspots and frontier trends of recommendation systems from the perspective of knowledge graph. *Expert Systems with Applications, 165*, Article 113764.
Shen, T., Mao, Y., He, P., Long, G., Trischler, A., & Chen, W. (2020). Exploiting structured knowledge in text via graph-guided representation learning. arXiv preprint arXiv:2004.14224.
Shi, C., Hu, B., Zhao, W. X., & Philip, S. Y. (2018). Heterogeneous information network embedding for recommendation. *IEEE Transactions on Knowledge and Data Engineering, 31*, 357–370.
Shi, C., Zhang, Z., Luo, P., Yu, P. S., Yue, Y., & Wu, B. (2015). Semantic path based personalized recommendation on weighted heterogeneous information networks. In *In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management* (pp. 453–462).
Simpson, E. H. (1949). Measurement of diversity. Nature, 163, 688–688.
Smith, F. A. (2007). Librarything. Reference Reviews.
Soares, L. B., FitzGerald, N., Ling, J., & Kwiatkowski, T. (2019). Matching the blanks: Distributional similarity for relation learning. arXiv preprint arXiv:1906.03158.
Song, W., Duan, Z., Yang, Z., Zhu, H., Zhang, M., & Tang, J. (2019a). Explainable knowledge graph-based recommendation via deep reinforcement learning. arXiv preprint arXiv:1906.09506.
Song, W., Xiao, Z., Wang, Y., Charlin, L., Zhang, M., & Tang, J. (2019b). Sessionbased social recommendation via dynamic graph attention networks. In *In Proceedings of the Twelfth ACM international conference on web search and data mining* (pp. 555–563).
Soylu, A., Corcho, O., Elvesæter, B., Badenes-Olmedo, C., Martı́nez, F. Y., Kovacic, M., Posinkovic, M., Makgill, I., Taggart, C., Simperl, E. et al. (2020). Enhancing public procurement in the european union through constructing and exploiting an integrated knowledge graph. In International Semantic Web Conference (pp. 430–446). Springer.
Steiner, T., Verborgh, R., Troncy, R., Gabarro, J., & Van de Walle, R. (2012). Adding realtime coverage to the google knowledge graph. In 11th International Semantic Web Conference (ISWC 2012) (pp. 65–68). Citeseer volume 914.
Sun, E., & Iyer, V. (2013). *Under the hood: The entities graph.* Facebook Engineering Blog.
Sun, S., Yang, D., Zhang, H., Chen, Y., Wei, C., Meng, X., & Hu, Y. (2018a). Important attribute identification in knowledge graph. arXiv preprint arXiv:1810.05320.
Sun, Y., Han, J., Yan, X., Yu, P. S., & Wu, T. (2011). Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. *Proceedings of the VLDB Endowment, 4*, 992–1003.
Sun, Z., Guo, Q., Yang, J., Fang, H., Guo, G., Zhang, J., et al. (2019). Research commentary on recommendations with side information: A survey and research directions. *Electronic Commerce Research and Applications, 37*, Article 100879.
Sun, Z., Hu, W., & Li, C. (2017). Cross-lingual entity alignment via joint attributepreserving embedding. In International Semantic Web Conference (pp. 628–644). Springer.
Sun, Z., Yang, J., Zhang, J., Bozzon, A., Huang, L.-K., & Xu, C. (2018b). Recurrent knowledge graph embedding for effective recommendation. In *In Proceedings of the 12th ACM Conference on Recommender Systems* (pp. 297–305).
Szekely, P., Knoblock, C. A., Slepicka, J., Philpot, A., Singh, A., Yin, C., et al. (2015). In *Building and using a knowledge graph to combat human trafficking* (pp. 205–221). Springer.
Tang, X., Wang, T., Yang, H., & Song, H. (2019). Akupm: Attention-enhanced knowledge-aware user preference model for recommendation. In *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1891–1899).
Tanon, T. P., Weikum, G., & Suchanek, F. (2020). Yago 4: A reason-able knowledge base. In European Semantic Web Conference (pp. 583–596). Springer.
Tousch, A.-M. (2019). How robust is movielens? a dataset analysis for recommender systems. arXiv preprint arXiv:1909.12799.
Uyar, A., & Aliyu, F. M. (2015). Evaluating search features of google knowledge graph and bing satori: Entity types, list searches and query interfaces. *Online Information Review*.
Van Loan, C. F. (2000). The ubiquitous kronecker product. *Journal of Computational and Applied Mathematics, 123*, 85–100.
Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledgebase. *Communications of the ACM, 57*, 78–85.
Wang, B., Huang, Y., & Li, X. (2016). Combining review text content and revieweritem rating matrix to predict review rating. Computational intelligence and neuroscience, 2016.
Wang, H., Wang, N., & Yeung, D.-Y. (2015). Collaborative deep learning for recommender systems. In *In Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1235–1244).
Wang, H., Wang, Z., Hu, S., Xu, X., Chen, S., & Tu, Z. (2019a). Duskg: A fine-grained knowledge graph for effective personalized service recommendation. *Future Generation Computer Systems, 100*, 600–617.
Wang, H., Zhang, F., Hou, M., Xie, X., Guo, M., & Liu, Q. (2018a). Shine: Signed heterogeneous information network embedding for sentiment link prediction. In *In Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 592–600).
Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X., et al. (2019b). Exploring high-order user preference on the knowledge graph for recommender systems. *ACM Transactions on Information Systems (TOIS), 37*, 1–26.
Wang, H., Zhang, F., Xie, X., & Guo, M. (2018b). Dkn: Deep knowledge-aware network for news recommendation. In *In Proceedings of the 2018 world wide web conference* (pp. 1835–1844).
Wang, H., Zhang, F., Zhang, M., Leskovec, J., Zhao, M., Li, W., et al. (2019c). Knowledge-aware graph neural networks with label smoothness regularization for recommender systems. In *In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 968–977).
Wang, H., Zhang, F., Zhao, M., Li, W., Xie, X., & Guo, M. (2019d). Multi-task feature learning for knowledge graph enhanced recommendation. In The World Wide Web Conference (pp. 2000–2010).
Wang, H., Zhao, M., Xie, X., Li, W., & Guo, M. (2019e). Knowledge graph convolutional networks for recommender systems. corr abs/1904.12575 (2019). arXiv preprint arXiv:1904.12575.
Wang, M., Liu, M., Liu, J., Wang, S., Long, G., & Qian, B. (2017). Safe medicine recommendation via medical knowledge graph embedding. ArXiv e-prints, (pp. arXiv–1710).
Wang, P., Fu, Y., Zhou, Y., Liu, K., Li, X., & Hua, K. A. (2020a). Exploiting mutual information for substructure-aware graph representation learning. In IJCAI (pp. 3415–3421).
Wang, R., Yan, Y., Wang, J., Jia, Y., Zhang, Y., Zhang, W., et al. (2018c). Acekg: A large-scale knowledge graph for academic data mining. In *In Proceedings of the 27th ACM international conference on information and knowledge management* (pp. 1487–1490).
Wang, X., He, X., Cao, Y., Liu, M., & Chua, T.-S. (2019f). Kgat: Knowledge graph attention network for recommendation. In *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 950–958).
Wang, X., Huang, T., Wang, D., Yuan, Y., Liu, Z., He, X., & Chua, T.-S. (2021). Learning intents behind interactions with knowledge graph for recommendation. In Proceedings of the Web Conference 2021 (pp. 878–887).
Wang, X., Wang, D., Xu, C., He, X., Cao, Y., & Chua, T.-S. (2019g). Explainable reasoning over knowledge graphs for recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence (pp. 5329–5336). volume 33.
Wang, X., Xu, Y., He, X., Cao, Y., Wang, M., & Chua, T.-S. (2020b). Reinforced negative sampling over knowledge graph for recommendation. In *In Proceedings of The Web Conference 2020* (pp. 99–109).
Wang, Y., Wang, M., & Fujita, H. (2020c). Word sense disambiguation: A comprehensive knowledge exploitation framework. *Knowledge-Based Systems, 190*, Article 105030.
Wang, Z., Zhang, J., Feng, J., & Chen, Z. (2014). Knowledge graph embedding by translating on hyperplanes. In Proceedings of the AAAI Conference on Artificial Intelligence. volume 28.
Wu, Y., & Wang, Z. (2018). Knowledge graph embedding with numeric attributes of entities. In *In Proceedings of The Third Workshop on Representation Learning for NLP* (pp. 132–136).
Xian, Y., Fu, Z., Muthukrishnan, S., De Melo, G., & Zhang, Y. (2019). Reinforcement knowledge graph reasoning for explainable recommendation. In *In Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval* (pp. 285–294).
Xiao, H., Huang, M., Hao, Y., & Zhu, X. (2015a). Transa: An adaptive approach for knowledge graph embedding. arXiv preprint arXiv:1509.05490.
Xiao, H., Huang, M., Hao, Y., & Zhu, X. (2015b). Transg: A generative mixture model for knowledge graph embedding. arXiv preprint arXiv:1509.05488.
Xiao, Y., Pei, Q., Yao, L., Yu, S., Bai, L., & Wang, X. (2020). An enhanced probabilistic fairness-aware group recommendation by incorporating social activeness. *Journal of Network and Computer Applications, 156*, Article 102579.
Xie, R., Liu, Z., Sun, M. et al. (2016). Representation learning of knowledge graphs with hierarchical types. In IJCAI (pp. 2965–2971).
Xin, X., He, X., Zhang, Y., Zhang, Y., & Jose, J. (2019). Relational collaborative filtering: Modeling multiple item relations for recommendation. In *In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 125–134).
Xu, B., Liang, J., Xie, C., Liang, B., Chen, L., & Xiao, Y. (2019). Cn-dbpedia2: An extraction and verification framework for enriching chinese encyclopedia knowledge base. *Data Intelligence, 1*, 271–288.
Xu, B., Xu, Y., Liang, J., Xie, C., Liang, B., Cui, W., & Xiao, Y. (2017). Cn-dbpedia: A never-ending chinese knowledge extraction system. In International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (pp. 428–438). Springer.
Yadav, A., & Vishwakarma, D. K. (2020). A unified framework of deep networks for genre classification using movie trailer. *Applied Soft Computing, 96*, Article 106624.
Yang, B., Yih, W.-t., He, X., Gao, J., & Deng, L. (2014). Embedding entities and relations for learning and inference in knowledge bases. arXiv preprint arXiv:1412.6575.
Yang, D., Guo, Z., Wang, Z., Jiang, J., Xiao, Y., & Wang, W. (2018). A knowledgeenhanced deep recommendation framework incorporating gan-based models. In 2018 IEEE International Conference on Data Mining (ICDM) (pp. 1368–1373). IEEE.
Yang, S., Liu, Y., Xu, Y., Miao, C., Wu, M., & Zhang, J. (2020). Contextualized graph attention network for recommendation with item knowledge graph. arXiv preprint arXiv:2004.11529.
Yang, Y., Esteban, C., & Tresp, V. (2016). Embedding mapping approaches for tensor factorization and knowledge graph modelling. In European Semantic Web Conference (pp. 199–213). Springer.
Yang, Z., & Dong, S. (2020). Hagerec: Hierarchical attention graph convolutional network incorporating knowledge graph for explainable recommendation. *Knowledge-Based Systems, 204*, Article 106194.
Ye, Y., Wang, X., Yao, J., Jia, K., Zhou, J., Xiao, Y., et al. (2019). Bayes embedding (bem) refining representation by integrating knowledge graphs and behavior-specific networks. In *In Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 679–688).
Yin, H., Cui, B., Li, J., Yao, J., & Chen, C. (2012). Challenging the long tail recommendation. arXiv preprint arXiv:1205.6700.

Yu, X., Ren, X., Gu, Q., Sun, Y., & Han, J. (2013a). Collaborative filtering with entity similarity regularization in heterogeneous information networks. *IJCAI HINA, 27*.

Yu, X., Ren, X., Sun, Y., Gu, Q., Sturt, B., Khandelwal, U., et al. (2014). Personalized entity recommendation: A heterogeneous information network approach. In *In Proceedings of the 7th ACM international conference on Web search and data mining* (pp. 283–292).

Yu, X., Ren, X., Sun, Y., Sturt, B., Khandelwal, U., Gu, Q., et al. (2013b). Recommendation in heterogeneous information networks with implicit user feedback. In *In Proceedings of the 7th ACM conference on Recommender systems* (pp. 347–350).

Yuan, J., Jin, Z., Guo, H., Jin, H., Zhang, X., Smith, T., et al. (2020). Constructing biomedical domain-specific knowledge graph with minimum supervision. *Knowledge and Information Systems, 62*, 317–336.

Zhang, C., Ŕe, C., Cafarella, M., De Sa, C., Ratner, A., Shin, J., Wang, F., & Wu, S. (2017). Deepdive: Declarative knowledge base construction. Communications of the ACM, 60, 93–102.

Zhang, F., Yuan, N. J., Lian, D., Xie, X., & Ma, W.-Y. (2016a). Collaborative knowledge base embedding for recommender systems. In *In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 353–362).

Zhang, J., Li, D., Li, P., & Huang, X. (2020a). Knowledge-graph-embedding-based question answering. US Patent App. 16/262,618.

Zhang, L., Liu, P., & Gulla, J. A. (2018a). A deep joint network for session-based news recommendations with contextual augmentation. In *In Proceedings of the 29th on Hypertext and Social Media* (pp. 201–209).

Zhang, L., Rettinger, A., & Zhang, J. (2016b). A probabilistic model for time-aware entity recommendation. In International Semantic Web Conference (pp. 598– 614). Springer.

Zhang, Q., Lu, J., Wu, D., & Zhang, G. (2018b). A cross-domain recommender system with kernel-induced knowledge transfer for overlapping entities. *IEEE Transactions on Neural networks and Learning Systems, 30*, 1998–2012.

Zhang, S., Wang, W., Ford, J., & Makedon, F. (2006). Learning from incomplete ratings using non-negative matrix factorization. In Proceedings of the 2006 SIAM international conference on data mining (pp. 549–553). SIAM.

Zhang, T., & Pan, Y. (2014). *Yelp challenge project report*. University of Washington Publication.

Zhang, Y., Ai, Q., Chen, X., & Wang, P. (2018c). Learning over knowledge-base embeddings for recommendation. arXiv preprint arXiv:1803.06540.

Zhang, Y., & Chen, X. (2018). Explainable recommendation: A survey and new perspectives. arXiv preprint arXiv:1804.11192.

Zhang, Y., & Wallace, B. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. arXiv preprint arXiv: 1510.03820.

Zhang, Y., Wang, J., & Luo, J. (2020b). Knowledge graph embedding based collaborative filtering. *IEEE Access, 8*, 134553–134562.

Zhang, Y., Zhang, M., Liu, Y., Ma, S., & Feng, S. (2013). Localized matrix factorization for recommendation based on matrix block diagonal forms. In *In Proceedings of the 22nd international conference on World Wide Web* (pp. 1511–1520).

Zhang, Z., Han, X., Liu, Z., Jiang, X., Sun, M., & Liu, Q. (2019). Ernie: Enhanced language representation with informative entities. arXiv preprint arXiv:1905.07129.

Zhao, C., Li, C., & Fu, C. (2019a). Cross-domain recommendation via preference propagation graphnet. In *In Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 2165–2168).

Zhao, C., Li, C., Xiao, R., Deng, H., & Sun, A. (2020). Catn: Cross-domain recommendation for cold-start users via aspect transfer network. In *In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 229–238).

Zhao, H., Yao, Q., Li, J., Song, Y., & Lee, D. L. (2017). Meta-graph based recommendation fusion over heterogeneous information networks. In *In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 635–644).

Zhao, J., Zhou, Z., Guan, Z., Zhao, W., Ning, W., Qiu, G., et al. (2019b). Intentgc: A scalable graph convolution framework fusing heterogeneous information for recommendation. In *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2347–2357).

Zhao, V. N., Moh, M., & Moh, T.-S. (2016). Contextual-aware hybrid recommender system for mixed cold-start problems in privacy protection. In 2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS) (pp. 400–405). IEEE.

Zheng, P., Xia, L., Li, C., Li, X., & Liu, B. (2021). Towards self-x cognitive manufacturing network: An industrial knowledge graph-based multi-agent reinforcement learning approach. *Journal of Manufacturing Systems, 61*, 16–26.

Zheng, W., & Zhang, M. (2019). Question answering over knowledge graphs via structural query patterns. arXiv preprint arXiv:1910.09760.

Zheng, Y., Tang, B., Ding, W., & Zhou, H. (2016). A neural autoregressive approach to collaborative filtering. In International Conference on Machine Learning (pp.764–773). PMLR.

Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., et al. (2016). Attentionbased bidirectional long short-term memory networks for relation classification. In *In Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers)* (pp. 207–212).

Zhou, Q., & Zhang, C. (2013). Relationship between scores and tags for chinese books—in the case of douban book. *Journal of Data and Information Science, 6*, 40.

Zhu, H., Chang, D., Xu, Z., Zhang, P., Li, X., He, J., Li, H., Xu, J., & Gai, K. (2019). Joint optimization of tree-based index and deep model for recommender systems. arXiv preprint arXiv:1902.07565.

Zhu, H., Li, X., Zhang, P., Li, G., He, J., Li, H., et al. (2018). Learning treebased deep model for recommender systems. In *In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1079–1088).

Ziegler, C.-N., McNee, S. M., Konstan, J. A., & Lausen, G. (2005). Improving recommendation lists through topic diversification. In *In Proceedings of the 14th international conference on World Wide Web* (pp. 22–32).

Zimmermann, A., Gravier, C., Subercaze, J., & Cruzille, Q. (2013). Nell2rdf: Read the web, and turn it into rdf. In KNOW@ LOD (pp. 2–8).

Further reading

Fellbaum, C. et al. (1998). Wordnet: An electronic lexical database mit press. *Cambridge, Massachusetts*.