Knowledge Graph-based Recommendation Systems: The State-of-the-art and Some Future Directions

Sajisha P.S., Anoop V.S*, Ansal K.A

Indian Institute of Information Technology and Management- Kerala (IIITM-K), Thiruvananthapuram, India - 695581
Data Engineering Lab, Indian Institute of Information Technology and Management- Kerala (IIITM-K), Thiruvananthapuram, India – 695581
Saintgits College of Engineering, Kottukulam Hills, Pathamuttam P. O, Kottyam, Kerala, India - 686532

Abstract

The unprecedented growth of unstructured data poses many challenges in semantic computing, which is an active research area for many years. While unearthing interesting patterns such as entities, relationships, and other metadata are important, it is equally important to represent them in an efficient, easy to access manner. Knowledge Graphs (KGs) are one such mechanism to represent facts extracted from unstructured text. KGs represent entities as nodes and relationships as edges. Such a representation may find applications in many meaning-aware computing applications such as question answering, summarization, etc., to name a few. Very recently, knowledge graph-based recommendation systems have become popular which has many advantages over traditional recommendation engines. This survey is an attempt to summarize and critically evaluate some of the very recent approaches to knowledge graph-based recommendation approaches.

1. Introduction

The astounding rate of generation of unstructured text data demands efficient algorithms and other methodologies to represent, analyze and extract useful patterns. Unearthing such patterns is crucial for businesses to generate actionable insights for key decision-making processes. Being a mechanism to represent entities and relationships using nodes and edges, knowledge graphs find applications in areas such as question answering, summarization, to name a few. When closely analyzing, it is found that knowledge graph-based recommendation systems are becoming popular and widely adopted, where recommendation systems are sub-classes of information filtering systems that suggest items to users based on certain features and conditions. These items can be as simple as books to read, films to watch and apparel to buy or even more complex suggestions in industries such as suggesting what to do next on the failure of a specific component. In the last few decades, we have witnessed a rise in web services including e-commerce portals such as Amazon, video portals such as YouTube and Netflix, etc. These platforms heavily use recommendation systems to suggest items to their users, which is unavoidable in our daily life.

Keywords

Knowledge Graphs, Recommendation Systems, Knowledge Representation, Semantic Computing, Machine Learning

* Corresponding Author
Anoop V.S.
Email: anoop.res15@iiitmk.ac.in
Most recommendation systems such as traditional collaborative filtering based approaches use simple user ratings done on items. These predefined information sources pose some constraints and thus the quality of the recommendation engines in many cases is degraded. While content-based and collaborative filtering based approaches are popular, there are approaches reported that use hybrid methodologies for recommendation tasks. Very recently, knowledge graph-based recommendation systems are found to be extremely useful when dealing with large sets of data. The ease in representing and suggesting contents makes the knowledge graph-based approaches very popular and we may find many production-level implementations nowadays. The advantage of graph-based recommendation systems is that heterogeneous data sources can be used for providing better recommendation results. Very recently, techniques using external knowledge bases for supplementing more contextual information are also introduced in the graph-based recommendation systems literature. Thus it is undoubtable that it is highly necessary to have systematic research in knowledge graph-based recommendation systems and this paper is an attempt in this direction.

This survey attempts to review some of the very recently reported approaches to knowledge graph-based recommendation algorithms. The major contributions of this survey are outlined as follows:

(a) Introduces knowledge graphs and some of the interesting areas where knowledge graphs may find applications.

(b) Critically review some of the prominent approaches that have been reported very recently in the literature that uses knowledge graphs for recommendation systems.

(c) Towards the end of this survey, the authors discuss some interesting research dimensions on knowledge graphs for recommendation algorithms.

Organization of this article: The remainder of this article is summarized as follows: Section 2 gives a brief introduction to semantic computing using knowledge graphs. Section 3 outlines a detailed survey on some of the very recent approaches for building recommendation systems using knowledge graphs. In Section 4, the authors highlight some of the future research trends in this dimension and Section 5 concludes the survey.

2. Knowledge Graphs and Semantic Web

With the seminal approaches to “semantic computing”, the researchers represent the knowledge using graphs consisting of nodes and edges. Such a representation mechanism has many advantages both in organizing and retrieving knowledge. The entities and concepts are represented as nodes and the relationships connecting these nodes are denoted by edges. For example, consider the sentence “A recommendation system is a subclass of information filtering system that seeks to predict the rating or preference”. In this case, entities such as “recommendation system”, “information filtering system”, “rating”, “preference” will be represented as nodes and possible relationships such as “subclass of”, “predict” will be represented as the edges connecting these entities.

The term “knowledge graph” was first introduced by Google in the year 2012 when they started to introduce the notion of semantics into their search algorithms. But now the term is widely used by other knowledgebase providers such as DBPedia. Anyhow, we consider any knowledge base as a knowledge graph if it exhibits (1) represents entities and their relationships in nodes and edges (2) the classes and/or concepts and their relationships in a schema or ontology (3) has wide coverage on various domains and (4) potentially link entities with other knowledge bases or graphs. There were several attempts reported in constructing knowledge graphs that resulted in various types of knowledge graph systems and tools introduced in this area. Some of such prominent systems are outlined in the subsequent sections.

(a) Cyc Knowledge Graph

Cyc Knowledge Graph [7] is one of the very early artificial intelligence projects that is treated as a long-living one that aims to integrate a comprehensive ontology and knowledge base. A miniature version of Cyc is also available, known as OpenCyc which is a curated knowledge graph with API end-points for the users to explore it. Cyc focuses on implicit knowledge that other AI platforms may use for their operations and enhancing knowledge [7].
(b) Freebase

Freebase is a publicly available crowdsourced method of creating knowledge graphs. It supports many of the real-world entity types in its schema but the organization maintaining Freebase, was acquired by Google and the Freebase was shut down in 2015. According to recent statistics, The last known version of Freebase contains nearly fifty million entities and three billion facts acquired over time.

(c) DBPedia

DBPedia is one of the most prominently used knowledge graphs that has wide coverage[8]. It is constructed by extracting structured information and facts available in Wikipedia. DBPedia contains millions of entities and their relationships that were primarily extracted from Wikipedia infoboxes. Text mining and natural language processing researchers and practitioners are heavily using DBPedia[8].

(d) Never Ending Language Learning (NELL)

NELL is a project by CMU[9] that attempts to solve the research question “Can computers read?” and from 2010, NELL is continuously running to extract facts from the unstructured text available in web pages. In addition to this, it also tries to improve its text understanding competence, so that it can extract more and more facts from the web, with high accuracy[9].

(e) Google Knowledge Graph

Introduced to the public in the year 2012, Google attempts to construct a large knowledge graph. Even though it is unclear how Google collects the facts and converts into a knowledge graph, practitioners often commented that they use resources such as Wikipedia to extract knowledge. Google’s knowledge graph contains billions of entities and their relationships and is still evolving.

3. Knowledge Graph-based Recommendation Systems

We use many applications on a daily basis where recommendation engines play a major role in recommending appropriate items to us based on our previous pattern of interaction with the system. For example, based on the articles you read, the recommendation engine can suggest other articles you are interested in or based on your purchase history, the items we are interested in, or music or playlist based on our previous interactions. A recommendation engine uses a large amount of data that has been generated by user interactions, analyzes it, extracts patterns and then personalized suggestions are made to satisfy the user's needs. While referring to the literature on recommendation systems, one can easily identify that Collaborative Filtering (CF) is one of the widely used and popular recommendation techniques. This approach makes use of the old interactions of a particular user and suggests or recommends contents based on their preferences. On the other hand, there are techniques that use comparison between items that have similar properties along with the past user interactions with the system for recommending items. This technique is called Content-Based Filtering (CBF). In a nutshell, CBF will recommend those items to the users, which have similar properties to those items the user has already bought or clicked. Researchers and practitioners of recommendation systems always attempt to device algorithms that leverage external knowledge for improving content-based recommendations. In a content-based recommendation engine, the features are normally available for both the users and the items. For users, these are typically their details such as age, location, etc., For items like books, the features may be the author, year of publication, name of the publisher, etc. In the recent past, there were several approaches proposed in the recommendation literature that uses content for recommendation, but a very few approaches have attempted the use of the interconnections between the content and external knowledge sources, which is referred to as a knowledge graph (KG).

Knowledge Graph is considered as a heterogeneous graph consisting of Nodes and Edges. Nodes correspond to entities and edges represent the corresponding relation. When comparing with approaches that do not use knowledge bases, the approaches that use KG into the recommendation has the following benefits[12]:
(1) A knowledge graph always exhibits a rich semantic relatedness among items that can be used to explore the hidden connections and thereby improve the accuracy of the recommendation results.

(2) The different types of relations in a KG helps to extend the user’s interest, which will increase the diversity of recommended items;

(3) KG connects items which are previously liked by a user which enables the explainability to the recommender systems.

In this direction, there are many recent approaches reported in the recommendation system literature and this survey attempts to critically evaluate and summarize some of those prominent approaches. Yixin Cao et al. recently proposed an approach [1] that unifies knowledge graph learning and recommendation. In this, the authors proposed an approach that has the capability of learning the model of recommendation and knowledge graph completion. The authors contributed a new recommendation model which is based on a translation-based approach, which considers several preferences in connecting a user to an item. The approach then trains it with a knowledge graph completion model by integrating different transfer schemes [1]. Their experiments on benchmarked knowledge graph datasets showed that their proposed approach significantly outperforms state-of-the-art knowledge graph-based recommendation systems. Another recent approach called KGAT [2], was proposed by Xiang Wang et. al., that models the connectivities which are high-order, in a knowledge graph, in a complete manner. It sends or transfers the embeddings from a node's neighbors recursively to refine the node's embedding. It also uses an attention mechanism that deals with discriminating the importance of the neighbors [2]. The authors have compared the performance of their proposed method with state-of-the-art approaches, such as Neural FM and RippleNet and found that the proposed approach significantly outperforms the state-of-the-art approaches [2].

A system that learns rules which are explainable for recommendation tasks, with knowledge graphs is reported very recently which was proposed by Weizhi Ma et. al.[3]. Majority of the previous approaches attempted to incorporate side information to get better recommendation results. But these approaches have their own disadvantages such as difficulties with explainability and debugging, requiring a heavy amount of human intervention and knowledge of the domain, to name a few. The authors claim that their proposed approach is an approach that uses a joint learning framework that induces explainable rules that are available in the knowledge graph that guides to building a neural recommendation model which is rule-guided[3]. Systematic experiments on a benchmarked dataset showcased the significant performance of the proposed approach in recommendation [3]. Another approach that uses a multi task feature learning framework for knowledge graph enabled recommendation is proposed by Hongwei Wang et. al. [4]. As collaborative filtering-based approaches usually exhibit issues such as sparsity and cold start problems with recommendations. Keeping these in mind, the authors proposed a multi-task feature learning approach (MKR) for knowledge graph enhanced recommendation. It is an end-to-end framework that makes use of KG embeddings to assist recommendation tasks[4]. The authors have demonstrated that their approach achieved better gains in real-world recommendation tasks such as recommendations in movies, books, music, and news recommendations, when compared with the prominent approaches in recommendation. The authors further claimed that MKR is successful to generate better recommendations even if it has to deal with sparse user-item interactions [4].

Another notable work reported was a knowledge-aware graph neural network based approach that incorporates label smoothness regularization for recommender systems [5]. To overcome the issue of manual feature engineering while developing knowledge graph-based recommendations, the authors proposed an approach that incorporated knowledge graph neural networks with label smoothness regularization for providing better recommendations for the users [5]. The authors first convert the knowledge graph into a weighted graph. Then they applied a graph neural network which is capable of computing item embeddings which preserves personalization features. A rigorous experiment has been conducted on four different datasets and the proposed system showed better performance compared to this method. A new approach called the Knowledge-aware Path Recurrent Network (KPRN) that uses KGs for recommendation was introduced by Xiang Wang et al. [6]. The advantage of KPRN is that it can provide path representations by combining the semantics entities and relationships. By considering the dependencies within a path which are sequential in nature, the authors conducted reasoning on these paths to leverage the available user-item interaction[6]. The authors have conducted experiments on movie and music datasets and demonstrated that their approach outperformed the state-of-the-art approaches in recommendations.
Very recently, an approach for learning knowledge graph embedding for products in e-commerce was reported by Da Xu et al. [10]. In this work, the authors proposed a novel knowledge graph embedding approach for learning the product relations as product knowledge for the electronic commerce domain [10]. They have constructed a new approach called a self-attention-enhanced distributed representation learning model that was capable of capturing embeddings from customer activity data from electronic commerce websites. The proposed approach was executed on some real-world dataset and the performance was evaluated for knowledge completion, search, ranking and recommendation processes [10]. Another approach that introduced a fine-grained knowledge graph for providing better personalized recommendations called DUSKG [11] was another approach that attempted incorporating multiple types of service data considering their logical relations. The authors proposed a compact data representation model incorporating different sorts of logical relations between the data. The five major relationships such as “FocusOn”, “BelongTo”, “USimilar”, “SSimilar” and “FSimilar” are considered for their experiment and it showed that their proposed approach showed better recommendation performance with least computation time [11].

Another recent approach that explored the higher order user preference on KGs for recommendation engines [12] was reported in the recommender system literature by Hogwei Wang et al. The authors stated that to avoid cold start problems, researchers normally use side information for building recommendation algorithms and in the proposed approach, the authors used a knowledge graph as the side information. They have proposed RippleNet which is a complete framework that uses a knowledge graph into the recommendation system [12]. The authors claim that this has brought more explainability in the recommendation engine and the real-world experiments showed that their proposed approach outperforms some of the state-of-the-art approaches for recommendation. A very notable and interesting work that explored the potential of using a knowledge graph specifically for movie recommendation, HI2Rec [13], was introduced by Ming He et al. Their proposed system used a mechanism that incorporated multiple information that is capable of learning the vector representations of users and items with the aim of providing top-N recommendations to the users [13]. The authors stated that this will solve the issue of earlier approaches which give less consideration about the properties or demographics of the users.

Zhu Sun et al. proposed an approach for learning KG embedding that can automatically learn meaningful representations of both entities and paths (relationships) between entities for representing an and taking into consideration, the preferences of the users towards items [20]. The proposed approach uses recurrent neural network architecture with a pooling operator. The duty of the pooling operator is to make sure the saliency of different edges in finding out the user preferences towards items. The authors have extensively validated the proposed approach on different real-world datasets which showed that their proposed approach significantly outperformed the state-of-the-art methods [20]. A deep network for news recommendation which makes use of the knowledge-aware feature was introduced by Hogwei Wang et al. [21] in which the authors address the problem of extended reasonability of the existing knowledge graph based-approaches. Their proposed approach, known as DKN, is a deep recommendation framework based on contents which can perform click-through rate predictions [21]. By conducting systematic experiments on real online news platforms, the authors were successful in showcasing the significant performance of their proposed approach [21]. A notable work that explored the possibility of using node2vec [22] for generating item recommendations by fusing machine learning with knowledge graph embeddings was introduced by Enrico Palumbo et al. [23]. The authors applied node2vec on a knowledge graph generated from the famous MovieLens 1M dataset and DBpedia and used the node relatedness to generate item recommendations. When compared with the traditional collaborative filtering-based approaches, the proposed method significantly outperformed them [23].

An approach called RippleNet [24] was reported in the recommendation systems literature recently. The system was developed keeping an aim of incorporating preferences of the users on knowledge graphs for recommendation systems. This approach stimulates the dissemination of preferences of users over a set of entities by extending a user’s potential interests along links in a KG [24]. Experiments conducted on benchmarked recommendation datasets show the RippleNet is capable of providing better recommendations when compared with the state-of-the-art approaches [24]. Most of the collaborative filtering based recommendation approaches only use the user-item rating matrix and do not consider semantic information. Ruihui Mu and Xiaqin Zeng proposed an approach based on a knowledge-graph to solve this issue [25]. They have used representation learning mechanisms on knowledge graphs to embed existing semantics into a vector space which is low-dimensional. Their proposed approach combines the semantics of items into the recommendation task by computing the meaning-aware similarity between items [25]. The authors have
claimed that their proposed approach could significantly outperform some of the state-of-the-art approach in knowledge graph-based recommendation engines.

There are other interesting approaches very recently reported in the knowledge graph based recommendation literatures such as unifying task oriented knowledge graph learning and recommendation[14], differentiated fashion recommendation using knowledge graph and data augmentation[15], a semantics driven knowledge graph for food recommendation[16], an attention-enhanced, knowledge-aware user preference model for recommendation[17], a graph augmented memory network for recommending medication combination[18], location embeddings for next trip recommendation[19], etc. which are worth mentioning in this paper.

4. Knowledge Graph-based Recommendation – Future Research Dimensions

Since technological advancements are still fueling and acting as a catalyst, the amount of unstructured text data that will be generated in the coming years will be highly exponential. Being an efficient knowledge representation mechanism, knowledge graphs will have huge potential in the coming years. Recommendation systems that need side information and augmented knowledge will heavily make use of knowledge graphs. This poses many challenges and research opportunities among text mining communities to devise better meaning-aware algorithms that can operate on knowledge-graphs for personalized recommendations. In this section, we discuss some of the potential research opportunities and avenues the recommendation system researchers and practitioners can look into.

One of the most interesting research dimensions on knowledge graph-based recommendation is to incorporate more side information into the recommendation engine so that better personalized recommendation is possible. For instance, bringing in more user information (say, demographic information) may improve the quality of recommendations. Also, connecting social network interactions would open new avenues to explore how social influence affects the recommendation. A very recent research trend in the recommendation system is the explainable recommendations. Integrating information propagation and decision processes, the explainability of recommendation engines may significantly improve, which is an area worth looking at. Graph Neural Networks (GCN) is one recent approach that got significant attention among recommendation system researchers and active research is happening in this dimension as well. Since earlier approaches face issues related to cold start problems in recommendation, it is worth experimenting on that dimension to eliminate the issues by bringing in techniques like zero-shot learning and reinforcement learning.

5. Conclusions

This survey outlined some of the very recent approaches in knowledge graph-based recommendation systems. As knowledge graph is one of the effective representation mechanisms for knowledge that has been unearthed from unstructured text, it got wider acceptance among research communities. A knowledge graph represents entities and relationships as nodes and edges respectively and a large number of meaning-aware applications and algorithms can operate on this graph. One such application is recommendation systems that suggest a user with items based on their previous interactions with the system. Knowledge graph based recommendation systems became very popular recently primarily due to its ability to supply side information for augmenting data and thus enhancing the quality of recommendations. This paper discusses some of the very prominent approaches reported very recently in the recommendation literature. Some interesting research dimensions are also discussed towards the end of this paper. This survey will be useful for the researchers and practitioners who wish to work on entity knowledge graphs based recommendation systems.
References


Author’s Biography

Sajisha P.S. is currently pursuing her research leaning to Master of Philosophy (M.Phil.) from Indian Institute of Information Technology and Management – Kerala (IIITM-K), Thiruvananthapuram, India. She has completed her Masters and Bachelors in Computer Science from Mahatma Gandhi University, Kottayam, Kerala. Her research interests include recommendation systems and knowledge graph-based machine learning applications.

Anoop V.S. has completed his Doctor of Philosophy (Ph.D) and Master of Philosophy (M.Phil) in Computer Science, both with specialization in Text Mining, Natural Language Processing, and Information Retrieval, from Indian Institute of Information Technology and Management - Kerala (IIITM-K), Thiruvananthapuram, India. Anoop is currently working as a Senior Scientist and Head of Training with Kerala Blockchain Academy (KBA) He is a passionate and active teaching professional with seven years of experience with Indira Gandhi National Open University, New Delhi. His major research interests include Applied Artificial Intelligence using Text Mining and Natural Language Processing (NLP) and Blockchains. He has several publications in his credit including international and national journals, book chapters and conference proceedings in major venues of Artificial Intelligence and Knowledge Management.
Ansal K A has completed his B.Tech in Electronics and Communication from Mahatma Gandhi University in 2007 with a University 3rd rank, and completed his post graduation in M.Tech Communications Systems from Anna University Gundy, Chennai with University Gold Medal. He has completed his Doctoral Degree from Pondicherry Central University in the area of Miniaturized Antennas with UGC JRF. He has an awardee of various fellowships like UGC MANF, DST Inspire, Research excellence award from IEAE, and young researcher award from IEI. He has more than 40 journals and 35 international conference publications in his credit. His area of interest belongs RF system design, Microwave and Millimetre wave, Wireless Communications, Digital image Processing, Machine and Deep learning. He has awarded International Engineering Educators Certification from IUCEE USA, and is an active member of IEEE, IEAE, IELISTE, IETE and IAENG.

How to Cite


doi: https://doi.org/10.30991/IJMLNCE.2019v03i03.004