Relation-aware Subgraph Graph Neural Network for Modeling Document Relevance

Zhenxiang Sun[1,2], Runyuan Sun[1,2] (✉), Zhifeng Liang[2], Bo Liu[1,2], Hui Liu[1,2], Zhenyu Li[1,2] and Hao Yuan[1,2]

[1] School of Information Science and Engineering, University of Jinan, 250022, China
sunry@ujn.edu.cn
[2] Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, 250022, China

Abstract. In the context of the information age, the exponential growth in the volume of document data makes it challenging to retrieve information quickly and accurately. Traditional keyword-based retrieval methods have limitations and cannot effectively capture the semantic information of a query, leading to irrelevant retrieval results. To improve the accuracy of retrieval, researchers have started to use knowledge graph (KG) tools to enhance the matching of document retrieval results, however, direct retrieval using graph structures is limited by exponential complexity and the inability to model distant related documents. To solve this problem, we propose a new information retrieval model, SGDR (Subgraph Neural Network-based Graph Representation and Document Retrieval), which utilizes relational subgraph neural networks to deeply mine the structural information and semantic associations in document KG. The SGDR models the relevance of documents mainly from semantic relations and local structure in the KG. The experimental results show that the SGDR model outperforms several baseline models on the DocIR dataset, including significant improvements in the key performance metric AUC. The effectiveness of each module in the model is verified through ablation experiments, and the results emphasize the importance of initializing a deep representation of the document knowledge graph.

Keywords: Government Entity Recognition, Multi-feature Fusion, Multi-headed Attention Mechanism, Semantic Representation.

1   Introduction

In the information age, where the amount of data is growing exponentially, it has become an important challenge to retrieve the required information quickly and accurately from the huge amount of data [1]. In the early stage of information retrieval, keyword-based retrieval methods are widely adopted for their simplicity and intuition. Users input keywords related to the query topic and the system returns documents containing these keywords. However, this approach shows obvious limitations when dealing with the complexity and diversity of natural languages [2]. It cannot effectively capture the semantic information of the query and often leads to irrelevant search results, the so-called "keyword bubble". As technology evolved, semantic-based retrieval methods emerged that attempted to improve retrieval accuracy by understanding the

semantic content of documents and queries. For example, search engines have begun to use natural language processing techniques to analyze the intent of a query and use metadata (e.g., titles, abstracts, keywords) in documents to assess relevance. Although these methods have improved retrieval quality to some extent, they still rely on explicit information in documents and are powerless to do anything about implicit, unarticulated knowledge. Knowledge graphs emerged as a structured knowledge representation that describes real-world entities and their interrelationships through the ternary form [1,2]. Earlier query methods based on knowledge graph directly use multi-hop query to retrieve relevant information on the constructed graph structure, it can only retrieve the proximal entities and for the distal related entities cannot be queried. In addition, the retrieval efficiency of multi-hop query in the graph structure is exponentially increasing (requiring breadth-first search), the retrieval efficiency is relatively low, and it is only applicable to sparse scenarios, which is less suitable for document retrieval and such high-density information management scenarios.

Knowledge Graph Representation Learning (KGRL), on the other hand, is a key technique in machine learning, such as graph machine learning [3,4,5] and knowledge graph embedding [6,7]. Based on this, we develop the Subgraph GNN for Document Retrieve (SGDR) algorithm based on subgraphs for document retrieval by taking advantage of the characteristics of a document knowledge graph that is centered on the document and diffuses around it according to the document's attributes for this application scenario of document retrieval in colleges and universities. SGDR models the document within the document subgraph range of document SGDR models document information in the scope of document subgraph, fully exploits the attributes, relationships, and associations between documents, and improves the modeling efficiency of document information. At the same time, it introduces the semantics of documents, models documents from both semantic and structural perspectives, provides comprehensive and detailed document characterization, and improves the accuracy of downstream retrieval.

In summary, the main contributions of this paper are:

- We innovatively model document relevance for document retrieval tasks from both semantic and structural perspectives, effectively supporting downstream document recommendation and intelligent search tasks.
- For modeling the semantic and structural relevance of documents, we designed a knowledge graph-based subgraph neural network model SGDR to model the relevance of documents more efficiently in the local structure.
- For both the document recommendation and document intelligent search tasks, SGDR shows superior performance to the baseline approach on our constructed document retrieval dataset.

2 Related work

2.1 Knowledge Graph Embedding

Knowledge Graph Embedding (KGE), which aims to map entities and relationships in the knowledge graph into a low-dimensional, continuous vector space. One of them,

TransE [15], is a knowledge graph embedding model based on translation transformations, proposed by Boris Bordes et al. in 2013. It stands for Translation-based Models (TBMs), which aim to capture structural and semantic information in a knowledge graph by mapping entities and relations into a low-dimensional vector space. However, it has some limitations, for example, it uses a linear assumption, which reduces the efficiency of modeling on graphs with complex relationships. Therefore, in order to enhance the modeling efficiency of embedded models, the TrasnH [10] model was proposed, TransH introduces the concept of hyperplanes, where entities are projected onto hyperplanes defined by relation vectors. In addition to the Trans family of translational transformation models spoken of above, another class of semantic-based models has been developed to apply more complex knowledge graph relationships and to fully exploit the semantics in the knowledge graph [16,17,18]. DistMult [3] introduces the concepts of distributed assumptions and multivariate relationships in the field of learning knowledge graph representations. To more effectively capture the complex interaction patterns between entities and relations to solve the complementation problem in knowledge graphs, RotatE is proposed. RotatE focuses on representing entities and relations in KGs as vectors and rotating the relation representations in the complex space to capture more complex semantic relations such as symmetry, and antisymmetric.

2.2 Graph Neural Network

Methods based on knowledge graph embedding are insufficient to mine the topology of the knowledge graph, so methods based on graph neural networks have emerged [20,21,22,23]. The unique advantages of graph neural networks (GNNs) in processing graph-structured data make them a powerful tool for knowledge graph modeling [24,25,26]. Graph Convolutional Network [11] (GCN) is a neural network architecture that learns on graph-structured data. GCNs can capture the feature information of nodes as well as topological relationships among nodes by performing convolutional operations on the nodes of a graph. This network architecture has shown excellent performance in tasks such as node classification, graph classification, link prediction, and knowledge graph completion, and can be effectively utilized for document retrieval. To model the weights of nodes during neighbor aggregation, Graph Attention Convolutional Network [12], or GAT for short, has been proposed, which learns the weights of neighboring nodes in a way that differentially models the information conveyed by messages. However, all these methods ignore the modeling of relationships in the graph, to model the semantic relationships in the knowledge graph, RGCN [13] assigns an independent weight matrix to each relationship type and performs differential modeling based on the corresponding weights of the relationships at the time of message aggregation. RGAT [14] proposes on the basis of RGCN to have weights computed for the neighbors under each relationship.

3 SGDR Model

3.1 Definitions

(1) Document Knowledge Graph

In this paper, we define an article knowledge graph $G = \{(h, r, t)|h, t \in E, r \in R\}$, where each ternary fact (h, r, t) describes the existence of a relation r between the head entity h and the tail entity t. For example, in a document knowledge graph, the ternary group (document 1, contains, Zhangsan), i.e., it denotes that the entity document 1 contains the relevant content of the entity Zhangsan, E denotes the set of all the entities, and R denotes all the set of relationships. Modeling these triples through the knowledge graph representation learning model can effectively capture the topological semantic relationships in the knowledge graph.

(2) Problem Definition

We consider the document retrieval task as a sorting problem, Specifically, **(i)** Recommended scenarios related to document recommendation: our goal is to associate documents based on user clicks or lookups. Given that the user is browsing the document doc, we first analyze the features of the document, to construct a feature representation of the document. Then, our model traverses the set of all possible candidate document entities $doc \in E_D$, performs feature extraction based on SGDR for each document entity, obtains the vectorized representation, and calculates its similarity with the document the user is browsing. **(ii)** Intent Matching and Document Similarity Modeling in Search Scenarios: We utilize the Bert language model to vectorize user query statements and model the semantic relationships between statements and related entities in the knowledge graph to facilitate further vector retrieval in the subsequent process, while a document-to-document similarity table, which records the similarity between each pair of documents in the set of candidate documents, can be utilized to facilitate the retrieval of the search structure. The construction of this similarity table is based on the semantic analysis of the document content, possibly using methods such as cosine similarity, Jaccard similarity, or semantic similarity measures based on deep learning.

3.2 Model overview

As shown in Figure 1, this paper proposes a model SGDR for document retrieval based on relational subgraph neural network mining of knowledge graph data semantics and subgraph structure for university document retrieval scenarios. The SGDR proposed in this paper mainly consists of three parts: (1) Initialization of knowledge graph. In this paper, for the constructed document knowledge graph, the knowledge graph representation learning model RotatE [8] is used to learn the representations of various attribute nodes in the document and the graph, which can effectively capture various complex relationships in the graph and improve the modeling of the structure and semantics of the document knowledge graph. (2) Relational subgraph extraction module. This module is mainly for the document knowledge graph to be modeled to do subgraph level extraction, the extracted relational subgraph contains a variety of complex attributes and entity information of the document, which can be effectively applied to the downstream relational graph neural network. (3) Relational subgraph neural network. This

module aims to model the relational associative relationships in the document subgraphs through the graph neural network model, and construct an overall representation of the subgraphs through the aggregation of the subgraph ranges, and the acquired overall representation of the subgraphs is subsequently applied to the information retrieval of the documents. In this paper, these three modules are used to construct a representation of a document from the document knowledge graph, and subsequently model the similarity between documents through the document representation and build a document similarity table for downstream information retrieval from users.
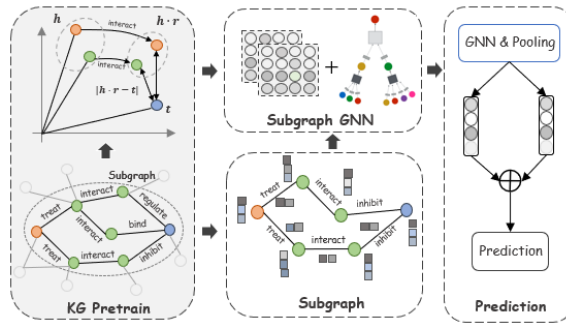

Fig. 1. Overview of the SGDR. SGDR contains three main modules: (1) The initialization of knowledge graph; (2) the extraction of document-based subgraph; (3) the subgraph graph neural network for representing extracted document-based subgraphs.

3.3 Initialization of Knowledge Graph

RotatE is a knowledge graph embedding method, which utilizes the structural properties of graphs and rotation operations to learn low-dimensional vector representations of entities and relations. The core idea of RotatE is to define relations over complex domains, such that relations can be represented by rotations of vectors to capture complex interactions between entities. In this paper, in order to effectively enhance the effect of relational subgraph neural network, we first initialize the entities and relations in the knowledge graph. Specifically, this paper adopts RotatE as the model for initializing the knowledge graph according to the characteristics of the constructed document knowledge graph. In the modeling process of knowledge graph, the characterization of entities and relations is one of the core issues. Given a ternary (h, r, t), we expect that we can have $X_t = X_h \cdot E_r$, where X_h and X_t denote the vector representations of the head entity and the tail entity, and E_r denotes the vector representation of the relation r. With the following optimization function, we can keep iterating to get the optimal vector representations of entities and relations:

$$s(h, r, t) = ||X_h \cdot E_r - X_t|| \tag{1}$$

where the symbol $\cdot$ denotes the multiplication between vectors. By minimizing $s(h, r, t)$ for positive samples and maximizing the fraction of negative samples, we can eventually obtain entity and relationship representations to initialize the knowledge graph.

3.4 Relational subgraph neural networks

(1) Document Subgraph Extraction

In this paper, we introduce an algorithm for mining document subgraphs based on the GraIL [9] model. Given a knowledge graph G_{kg} and a document pair (u, v), this paper extracts the k-order neighbors of the node based on the head node u to obtain its neighbor node $N_k (u)$, and at the same time extracts the k-order neighbors of the node based on the head node v to obtain its neighbor node $N_k (v)$, and subsequently obtains the intersection of the two neighbors based on $N_k (u)$ and $N_k (v)$ obtained from the knowledge graph. Thus, the common nodes of the two nodes are obtained and finally a PATH is formed. We show the schematic diagram of how to obtain a path by multi-hop search in Figure 3. Finally, extract the edges of these nodes with each other from the knowledge graph to form a subgraph $g = (V, E)$.

(2) Subgraph Neural Networks

This paper designs a subgraph neural network, the subgraph extracted from the document knowledge graph contains associations between documents. To fully mine the effective semantic information in the knowledge graph, we design a relational subgraph neural network, we initialize the representation of the relational subgraph neural network based relationship embedding of RotatE, as specified in Eq (2):
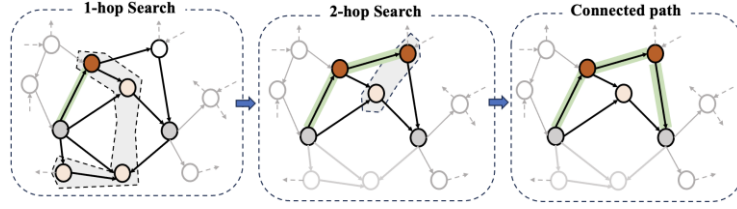

Fig. 2. Individual path acquisition process in document subgraph extraction.

$$\mathbf{x}_i^l = \sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}_r(i)} \alpha_{i,r} \mathbf{W}_r^l \phi\big(\mathbf{e}_r^{l-1}, \mathbf{x}_j^{l-1}\big),$$

$$\alpha_{i,r} = \text{sigmoid}\big(\mathbf{W}_1\big[\mathbf{x}_i^{l-1} \oplus \mathbf{x}_j^{l-1} \oplus \mathbf{e}_r^{l-1}\big]\big), \tag{2}$$

where W_r^l denotes the parameters of the neural network of layer l on the relation r, $e_r^{(l-1)}$ denotes the feature representation of the edges in the subgraph of layer $l - 1$ on the relation r, X_i^l denotes the vector representation of the ith node of layer l, $\phi(\cdot, \cdot)$ denotes the feature aggregation function, which is used to transform the features of the nodes and the edges and aggregated to the target node, and $\oplus$ denotes the feature vectors between the $\oplus$ denotes the splicing operation between the feature vectors, which is used to aggregate the information of nodes and edges. In this way, we can effectively learn the vector feature representation of nodes and relations, thus improving the modeling of the whole graph.

3.5 Information retrieval and optimization

(1) Recommended Documents for Recommended Scenarios

For the vector feature representation of entities and relationships learned by the subgraph neural network encoding, we introduce an optimization objective of modeling document relevance to update the iterative subgraph neural network model and the representation of the knowledge graph. Specifically, to effectively optimize this objective, we constructed an optimization dataset based on the correlation between documents. This dataset contains document pairs with relatively high correlation as positive samples and document pairs with relatively low correlation as negative samples. Optimization is subsequently performed on this dataset. Given a pair of documents (u, v), we obtain the representations of document u and document v from the document vector representations X obtained from the subgraph neural network, i.e., $h_u = X[:, u]$ and $h_v = X[:, v]$, and then we merge the obtained document representations and transform the features, and get the probability of correlation between the pairs of documents after the sigmoid activation function, and then calculate the probability of correlation and true label between the document pairs. Then we calculate the error between the correlation probability and the real label between document pairs and optimize the vector representation of subgraph neural network and knowledge graph by this error. The details are shown in Eq (3):

$$p_{(u,v)} = \sigma\big(f([\mathbf{h}_u \oplus \mathbf{h}_v])\big), \tag{3}$$

where $\mathbf{h}_u$ and $\mathbf{h}_v$ are vector representations of document nodes u and v, $f(\cdot)$ is a feature transformation function, $\sigma(\cdot)$ is a sigmoid activation function to map the association between documents into a probability space, and $p_{(u,v)}$ denotes the probability of correlation between the pair of documents. Subsequently, in order to realize the classification of positive and negative samples of document association, this paper introduces the cross-entropy loss for this objective:

$$\ell(u, v) = -\sum_{r \in \mathcal{R}} \log\big(p_{(u,v)}\big) y_{(u,v)}, \tag{4}$$

where $p_{(u,v)}$ is the association probability of the sample pair (u, v) and $y_{(u,v)}$ is the true association of the document pair (1 for positive samples and 0 for negative samples). By optimizing the subgraph neural network and the probabilistic scoring function for information retrieval with this loss function, a document-to-document association can be modeled and the final vector representation of the document obtained can be effectively used in relevant document recommendation scenarios.

(2) Intent Matching and Document Similarity Modeling in Search Scenarios

In the application scenario of document management in colleges and universities, search is a very active behavior of the user, which fully expresses the user's intention. Among them, it is especially important to effectively model the user's search intent and establish the association between it and the document knowledge graph, which is the key to realize intelligent search. In this paper, we rely on the subgraph neural network model proposed above, and introduce a module for user search intent understanding, i.e., semantic modeling of user search utterances through the Bert language model, and joint optimization of the Bert language model and SGDR by constructing the associated data with entities in the knowledge graph. Specifically, to address this goal, we

construct a semantically and entity-associated dataset. This dataset contains association pairs between problematic utterances and documents, and contains positive and negative samples of associations between utterances and articles. Given an utterance and document pair (s, u), we use Bert to model the vectorized representation h_s of the utterance, while using the subgraph neural network SGDR model to characterize h_u of the document u in the Knowledge Graph. We then perform a feature transformation for the two representations and compute the probability of the association between the two using a sigmoid function as shown in Eq(5):

$$p_{(s,u)} = \sigma\big(f([\mathbf{h}_s \oplus \mathbf{h}_u])\big), \tag{5}$$

where h_s and h_u are vector representations of document node u and question utterance s, $f(\cdot)$ is a feature transformation function, $\sigma(\cdot)$ is a sigmoid activation function to map the association between documents and utterances to the probability space, and $p_{(s,u)}$ denotes the probability of correlation between documents and utterances. Eventually, the correlation probability of the problem utterance and document is calculated with the error of the real label, for which the subgraph neural network SGDR, the semantic understanding model BERT, and the representation of the knowledge graph are optimized. The details are shown in Eq(6):

$$\ell(s, u) = -\sum_{s \in S} \log\big(p_{(s,u)}\big) y_{(s,u)}, \tag{6}$$

where $y_{(s,u)}$ is the true association between them (1 for positive samples and 0 for negative samples.) From this we obtain a document retrieval model that can be used in search scenarios.

4   Experiments

4.1   Dataset

We will introduce the datasets to be used in this paper from two perspectives.

(1) Document correlation dataset

For each document entity, can be associated with a single attribute node as a single document as a strongly associated document, it is constructed as a positive sample. At the same time for each document, by randomly selecting ten times the number of documents compared to the positive samples in all the document entity candidate sets to construct negative samples, we can finally get the positive and negative samples 1:1 data set DocIR (Document information retrieve). DocIR contains 1,524 pairs of positive samples, 15,240 pairs of negative samples constructed randomly, and the total number of samples is 16,764. Optimizing the information retrieval task with this dataset can effectively improve the effectiveness of related document retrieval.

(2) Q&A and document correlation dataset

We propose a question statement and document relevance dataset DocSD (document search dataset), which contains a batch of correlation samples between query statements and documents. For each attribute, we construct templates such as "Find {:attribute value} related documents", "Please help me retrieve the documents of {:attribute value} department. " and so on templates, for each document to construct 10 query statements,

so that for each document to build 10 query statements and documents between the positive samples. We randomly sampled 20 statements constructed from other documents for constructing negative sample pairs between and documents. Finally, in the DocSD dataset, we constructed 3810 positive samples and 7620 negative samples.

4.2   Experimental settings

(1) Hyperparameter Settings.
We used a grid search method to optimally retrieve the hyperparameters used by the model, specifically, we searched from 32-256 with an embedding dimension of 128 dimensions, from 1-5 with a negative sampling number of 4, with a learning rate of 0.05, and with the optimizer set to Adam.
(2) Baselines
We consider the following models as our baselines:
TransE [10]: an early knowledge graph embedding model. The core idea of TransE is to map entities and relations into the same vector space and assume that for any given fact, **RotatE** [8]: is to operate on complex domains to better capture the symmetry and anti-symmetry relationships between entities. **GCN** [11] is a neural network model that performs convolutional operations directly on graphs. **GAT** [12] is a model that introduces an attention mechanism on top of GCN. **RGCN** [13] is a neural network model specialized for processing graphs with multiple types of relationships. **RGAT** [14] is a combination of RGCN and GAT.

5   Results and Discussion

5.1   Overall performance

We introduce the widely used evaluation metrics Hits@1, Hits@10 and MRR as shown in Table 2.

Table 2. Experimental results on Document retrieval KG.

Models	Hits@1	Hits@10	MRR
TransE	0.3682	0.6541	0.4667
RotatE	0.4674	0.7789	0.5129
SGDR	**0.4980**	**0.7922**	**0.5679**

Based on the observation of Table 3, it can be known that SGDR improves the ranking effect Hits@1, Hits@10 and MRR of knowledge graph characterization by 3.06%, 1.33% and 5.5%, respectively, compared to the baseline model. This shows that our designed SGDR can enhance the comprehensive characterization of document knowledge by further improving the RotatE model, which enhances the modeling of document information from the effect of data.

On the other hand, in this paper, we construct a model for document retrieval for the graph neural network approach and test it on our constructed dataset DocIR, and the test results are shown in Table 3:

Table 3. Experimental results on DocIR.

Models	ACC	AUC	AUPR
GCN	0.7937	0.8341	0.8209
GAT	0.8009	0.8301	0.8322
RGCN	0.8329	0.8577	0.8801
RGAT	0.8298	0.8512	0.8736
SGDR	**0.8547**	**0.8681**	**0.8976**

The results presented in Table 3 reveal the potential and advantages of Relational Graph Neural Networks (RGNNs). The core strength of relational graph neural networks lies in their ability to capture and utilize the diverse patterns of relationships among entities, which is crucial for improving the effectiveness of document retrieval tasks. SGDR achieves significant improvements: accuracy (ACC), area under the curve (AUC) and area under the precision-recall curve (AUPR). This indicates that SGDR not only outperforms the traditional baseline model in terms of overall performance, but also shows a stronger ability to recognize positive class samples when dealing with unbalanced datasets through the improvement in AUPR metrics. Meanwhile, in order to verify the effectiveness of the model SGDR in intelligent search scenarios, we tested it on the dataset DocSD and the results are shown in Table 4.

These improvements in the SGDR may stem from its unique network architecture, optimized hyperparameter settings, or more efficient feature representation learning methods. Through well-designed message passing and attention mechanisms, SGDR is able to capture the correlation between documents and queries more accurately, thus achieving higher performance in retrieval tasks.

Table 4. Experimental results of the UATD.

Models	ACC	AUC	AUPR
GCN	0.8266	0.8421	0.8343
GAT	0.8031	0.8459	0.8565
RGCN	0.8421	0.8612	0.8904
RGAT	0.8324	0.8693	0.8532
SGDR	**0.8856**	**0.8977**	**0.9065**

5.2 Ablation Study

We conducted ablation experiments: (1) **SGDR w/o IR**: This model removes the representation initialization module of the knowledge graph for the original SGDR model and uses random initialization for subsequent experiments. (2) **SGDR w/o SE**: verify the effectiveness of the module by removing the subgraph extraction module. (3) **SGDR w/o SG**: This model removes the relational subgraph neural network module.

For these three variant models of the original model SGDR, we tested their AUC, AUPR on the dataset DocIR, respectively. The SGDR model reaches optimization when all three modules are included at the same time, and there is a substantial decrease in the effect after removing all three modules separately, which shows the effectiveness of each module in our design.
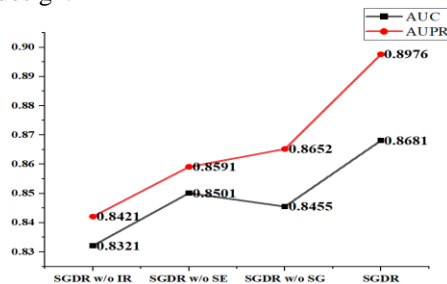

Fig. 3. The results on DocIR dataset

6   Conclusion

We propose SGDR based on knowledge graph representation learning to address the shortcomings of traditional keyword retrieval methods in the field of information retrieval, which effectively improves the accuracy and relevance of document retrieval by combining the rich semantic information of knowledge graphs and the representation learning technology. The core of SGDR model is to use relational subgraph neural network to deeply mine the structural and semantic information of document knowledge graph, and construct a fine-grained representation of documents through three steps of initializing knowledge graph representation, subgraph extraction and subgraph neural network modeling, and build a document similarity table on the basis of which to realize efficient information retrieval.

References

1. Landolsi, Mohamed Yassine, Lobna Hlaoua, and Lotfi Ben Romdhane. "Information extraction from electronic medical documents: state of the art and future research directions." Knowledge and Information Systems 65.2 (2023): 463-516.
2. Panja, Srijit. "Information Retrieval Systems in Healthcare: Understanding Medical Data Through Text Analysis." Transformative Approaches to Patient Literacy and Healthcare Innovation. IGI Global, 2024. 180-200.
3. Zhang, Ziwei, Peng Cui, and Wenwu Zhu. "Deep learning on graphs: A survey." IEEE Transactions on Knowledge and Data Engineering 34.1 (2020): 249-270.
4. Xia, Feng, et al. "Graph learning: A survey." IEEE Transactions on Artificial Intelligence 2.2 (2021): 109-127.
5. Qiao, Lishan, et al. "Data-driven graph construction and graph learning: A review." Neurocomputing 312 (2018): 336-351.
6. Zou, Xiaohan. "A survey on application of knowledge graph." Journal of Physics: Conference Series. Vol. 1487. No. 1. IOP Publishing, 2020.

7. Wang, Quan, et al. "Knowledge graph embedding: A survey of approaches and applications." IEEE transactions on knowledge and data engineering 29.12 (2017): 2724-2743.

8. Sun, Zhiqing, et al. "Rotate: Knowledge graph embedding by relational rotation in complex space." ICLR (2019).

9. Teru, Komal, Etienne Denis, and Will Hamilton. "Inductive relation prediction by subgraph reasoning." International Conference on Machine Learning. PMLR, 2020.

10. Wang, Zhen, et al. "Knowledge graph embedding by translating on hyperplanes." Proceedings of the AAAI conference on artificial intelligence. Vol. 28. No. 1. 2014.

11. Kipf, Thomas N., and Max Welling. "Semi-supervised classification with graph convolutional networks." arXiv preprint arXiv:1609.02907 (2016).

12. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. stat, 1050(20), 10-48550.

13. Schlichtkrull, Michael, et al. "Modeling relational data with graph convolutional networks." The semantic web: 15th international conference, ESWC 2018, Heraklion, Crete, Greece, June 3–7, 2018, proceedings 15. Springer International Publishing, 2018.

14. Ishiwatari, Taichi, et al. "Relation-aware graph attention networks with relational position encodings for emotion recognition in conversations." Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP). 2020.

15. Fan, Miao, et al. "Transition-based knowledge graph embedding with relational mapping properties." Proceedings of the 28th Pacific Asia conference on language, information and computation. Waseda University, 2014.

16. Ji, Guoliang, et al. "Knowledge graph embedding via dynamic mapping matrix." 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). 2015.

17. Guo, Shu, et al. "Semantically smooth knowledge graph embedding." 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). 2015.

18. Lin, Yankai, et al. "Learning entity and relation embeddings for knowledge graph completion." Proceedings of the AAAI conference on artificial intelligence. Vol. 29. No. 1. 2015.

19. Zhang, Zhao, et al. "Knowledge graph embedding with hierarchical relation structure." Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018.

20. Chen, Ming, et al. "Simple and deep graph convolutional networks." International conference on machine learning. PMLR, 2020.

21. Yang, Yiding, et al. "Factorizable graph convolutional networks." Advances in Neural Information Processing Systems 33 (2020): 20286-20296.

22. Mao, Kelong, et al. "UltraGCN: ultra simplification of graph convolutional networks for recommendation." Proceedings of the 30th ACM international conference on information & knowledge management. 2021.

23. Hong, Danfeng, et al. "Graph convolutional networks for hyperspectral image classification." IEEE Transactions on Geoscience and Remote Sensing 59.7 (2020): 5966-5978.

24. Yang, Yiding, et al. "Distilling knowledge from graph convolutional networks." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

25. Yu, Donghan, et al. "Knowledge embedding based graph convolutional network." Proceedings of the web conference 2021. 2021.

26. Song, Xiangyu, et al. "Jkt: A joint graph convolutional network based deep knowledge tracing." Information Sciences 580 (2021): 510-523.