EITCMQA: MULTI-HOP QUESTION ANSWERING METHOD OF TCM KNOWLEDGE GRAPH BASED ON CONVOLUTIONAL NEURAL NETWORK

Xiaohui CUI[1], Yuze SUN[2], Yunkai SHEN[3], Dongmei LI[4,*], Xiaolong FENG[5], Changzhen SHI[6], Xiaoping ZHANG[7,*]

With the widespread adoption of artificial intelligence and advancements in traditional Chinese medicine (TCM) information, numerous TCM online consultation platforms have emerged. These platforms allow users to enter their symptoms online and receive corresponding prescriptions automatically. However, many of these question-answering (QA) systems lack research on text semantic relevance, leading to poor model performance. To address these issues, we propose a QA model called EITCMQA that integrates multi-hop reasoning and convolutional neural networks, while also constructing a TCM knowledge graph. Firstly, we construct the knowledge graph by utilizing desensitized data from prestigious Chinese physicians at the China Academy of Chinese Medical Sciences and open-source knowledge graph data. We then embed the entities and relationships within it. Users' questions are represented by vectors, and finally, the multi-hop knowledge graph reasoning method score and link prediction score are combined to enhance the multi-hop TCM knowledge question answering approach. Our experiments were conducted on the cMedQA2 dataset, a representative medical field question answering dataset, and the results indicate that our model outperforms existing

[1] School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China
[2] School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China
[3] School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China
[4] Prof., School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China; email: lidongmei@bjfu.edu.cn
[5] School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China
[6] School of Information Science and Technology, Beijing Forestry University, Beijing, China; National Forestry and Grassland Administration Engineering Research Center for Forestry oriented Intelligent Information Processing, Beijing, China
[7] Prof., National Data Center of Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing 100700, China; email: xiao_ping_zhang@139.com

models. EITCMQA improves the adaptability to complex networks, increases the speed and accuracy of computation, and enhances the dimension of answer scoring.

Keywords: knowledge graph, question answering, multi-hop reasoning, link prediction, traditional Chinese medicine

1. Introduction

Knowledge Graph (KG) is a method for representing and storing knowledge that involves recording entities and concepts in the real world, as well as the relationships between them. With the increasing popularity of artificial intelligence and the development of Traditional Chinese Medicine (TCM) informatization, TCM knowledge has been transformed into entities and relationships within KGs for structured storage. Medical science has become one of the most widely used vertical fields for KGs. The Institute of Traditional Chinese Medicine Information, Chinese Academy of Sciences of Traditional Chinese Medicine has constructed nine knowledge graphs, such as the TCM language system, the TCM prescription knowledge graph, and the TCM cosmetology knowledge graph. These knowledge graphs are embedded in the knowledge service platform of traditional Chinese medicine, TCMKB, and are visually displayed. This makes it convenient for patients and other users to access domain knowledge at the conceptual level, allowing for the application of knowledge retrieval, question answering, and decision making [1].

TCM KG provides a methodological foundation for knowledge question answering and can be used to facilitate diagnosis and treatment in the era of artificial intelligence. By understanding patients' symptoms and questions based on the TCM KG, answers such as corresponding drug prescriptions can be inferred or found from the knowledge graph. Currently, mainstream QA methods based on KGs are divided into two categories. Semantic analysis-based methods typically require manual annotation of "question-logic expression" data, which can be costly. As the TCM knowledge graph continues to expand rapidly, development using these methods is limited. Answer-ranking methods use a ranking model to score different answer candidates and select the set of answer candidates with the highest score as the answer corresponding to the question. Cao et al. [2] used the answer-ranking method in knowledge retrieval question and answer based on the liver cancer medical knowledge graph. However, these two methods rely on the graph being relatively comprehensive and complete. When the KG is incomplete, the accuracy of QA will be limited. Link prediction involves predicting the relationship between head and tail entities to address incompleteness in the KG. It is widely used in QA based on TCM KG to solve this problem. For example, Yang et al. [3] proposed an innovative inductive reasoning model called Hierarchical Structure Type (HSTP). The HSTP model obtains information based on types and hierarchical structures and uses the semantic association between entities in the KG to integrate types, nodes, and relations. Next, they proposed a relational correlation network to learn the importance of different patterns for inductive link prediction and finally obtained a real and complete TCM medical case KG. These link prediction methods focus on shallow, fast models that can be extended to large-scale KGs. However, external models learn fewer features than deep models, which can limit performance.

For example, in a medical QA system, when the question "Headache of blood stasis syndrome" appears, many different entities may be selected based on the deep learning model using KGs, resulting in multiple entity pairs. After scoring and sorting the candidate entities, the highest score is used as the output answer. However, this approach may overlook situations where there are multiple answer entities or where there are multiple paths between the same entity pair. In the TCM KG, for instance, there are two triples with different tail entities: 'blood vessel → clogging → headache' and 'blood vessel → blocking → angina pectoris'. Additionally, there are two triples with different paths: 'blood vessel → blockage → headache' and 'blood vessel → relaxation → headache'. Although both pathways express the same symptoms, they have distinct causes. Different paths in KG can reflect varying relationships between TCM head and tail entities. However, some of these paths may be inconsistent with the context of the question-answer and inconsistent with doctors' reasoning logic. Therefore, selecting the most suitable path and optimal answer is critical in correctly answering patients' symptoms.

To address the challenges mentioned above, the current mainstream QA scoring mechanism based on knowledge graph embedding is relatively limited. After scoring and sorting candidate entities, only the one with the highest score is used as the answer output, which may result in multiple candidate answers being overlooked. To solve this problem, we propose an improved multi-hop TCM knowledge graph question answering method called EITCMQA, which is based on a convolutional neural network. This method combines the link scores obtained through multi-hop analysis with the link prediction scores obtained through convolutional neural networks. By doing so, the best possible physical answers can be selected from all candidate entities sharing the same relation link. This effectively solves the problem of missing answers and improves the adaptability of complex networks, speed, and calculation accuracy.

Our contributions are summarized as follows:

(1) We construct a knowledge graph in the field of TCM, using the diagnostic data of prestigious Chinese physician from China academy of Chinese medical sciences and the open-source medical KG as data support.

(2) We propose a novel TCM knowledge question answering method, EITCMQA, which combines the convolutional neural network with the multi-hop question answering inference model to improve its performance in TCM knowledge graph question answering.

(3) Experiments on the QA dataset in the medical field demonstrate the effectiveness of IETCMQA on multi-hop knowledge question answering.

2. Paper contents

2.1 Related Work

2.1.1 TCM knowledge graph. Currently, numerous TCM scholars have built and implemented TCM KG in various subfields by utilizing contemporary literature and data sourced from reliable websites. Odema et al. [4] developed CMeKG 1.0, the first edition of the Chinese medical knowledge graph (KG), containing over 100,000 instances of medical concept relations. Zhou et al. [5] built a TCM syndrome KG using modern literature on syndrome differentiation, such as TCM syndrome ontology and Chinese encyclopedia websites. Using this KG, they created a TCM health management platform and an intelligent question answering model. Yin et al. [6] constructed the KG of prescriptions based on the knowledge of classical famous prescriptions or prescriptions in the prescriptions database and applied it to applications such as retrieval, visualization, and graph pattern search. Zhang et al. [7] created TCM core knowledge graph, TCM knowledge graph, and prescription experience knowledge graph by combining information from various data sources. Liu et al. [8] constructed a knowledge graph based on Neo4j graph database using Yao Naili's clinical experience as the research object, enabling visual display and semantic search functions. Their system provides a new tool for analyzing and displaying physicians' clinical thinking and diagnosis and treatment characteristics.

2.2.2 Multi-hop question answering method based on medical KG. Bordes et al. [9] suggested using low-dimensional embedding vectors to represent questions and entities and calculating vector similarity to evaluate potential answers. Saxena et al. [10] applied the knowledge graph embedding method to generate relation vector representations in the knowledge graph embedding and used the scoring function of the knowledge graph embedding model to score candidate answers, with the help of a sentence vector embedding model. It effectively alleviates the performance loss caused by the sparsity of KG. ComplEx [11] represents entity vectors and relation matrixs in the ComplEx space. ConvE [12] is an embedding method based on CNN, which applies convolution filters to 2D reconstruction of entity and relation embedding to capture rich interactions among its components. C Shang et al. [13] used graph network as encoder and Conv-TransE as decoder for knowledge graph completion task. RotatE [14] projects entities in a ComplEx space, and relations are represented as rotations on ComplEx plane. InteractE [15] increases the interaction among vectors of entities and relations and improves link prediction accuracy by feature replacement, reshaping operations, and cyclic convolution. TransferNet [16] supports both label and text relations for multi-hop QA. Liu Y et al. [17] innovated the Polo method based on a biomedical KG, which combined reinforcement learning strategies with logical rules and applied it to bioinformatics databases.

Sun et al. [18] proposed TCM auxiliary diagnosis and treatment system based on artificial intelligence for rheumatoid arthritis. The system utilizes patient medical records and joint image data to accurately diagnose rheumatoid arthritis and associated syndromes, aiding doctors in prescribing TCM treatment based on individualized syndromes. Liu et al. [19] first constructed the KG in the medical field and encoded the knowledge through the knowledge representation model. Combined with the patient's complaint text, the representation vector of the patient's symptom entity was obtained. Then the patient's complaint representation vector and the index representation vector were used to assist in the diagnosis of complications through CNN-DNN network. Gu et al. [20] developed a stroke disease TCM syndrome classification model using support vector machine, with age, gender, height, weight, and common clinical symptoms of stroke including main symptoms as input and common clinical syndrome judgment as output.

2.2 Methods

The TCM knowledge graph data constructed in this paper comes from two parts, one of which is the medical records of famous TCM veterans from China Academy of Chinese Medical Sciences. The data of chief complaint, history of present disease, symptoms, laboratory examination, physical examination, medical diagnosis, traditional Chinese medicine prescription, western medicine and so on were covered after the desensitization process. We process and normalize the data, obtain entity relations in the form of triples, and save them as JSON data files that can be directly imported. The data in JSON is directly imported into the Neo4j graph database. The graph contains 7499 entities and 12,520 relationships. Since most of this data is normalized, we can directly import the data into the Neo4J graph database using CSV format files, without any extra processing. And the database is not neat basis, according to the part of the old Chinese medicine basis. We first use the problem template to segment the medical record prescriptions into separate prescriptions for each condition. Then the TF-IDF segmentation method was used to refine the medical records, and this part of data was added to the knowledge graph according to the attribute principle.

We use DiseaseKG on OpenKG: Based on the cnSchma common disease information knowledge graph as another part of the knowledge graph database. The graph contains 44656 entities, and 312159 relationships, covering several entity types such as diseases, drugs, food, inspection items, and subjects. The number of specific entities and relationships of the two types of knowledge graphs is shown in Table 1.

Table 1

Relevant entity relation data

Data source	Entity	Relation
TCM knowledge graph constructed by the Chinese Academy of TCM	7499	12520
DiseaseKG on OpenKG: Based on the cnSchma common disease information knowledge graph	44656	312159
Total	52155	324679

Our goal is to calculate the score of all the candidate answers and sort them to select the best answer to the patient's problem. Fig. 1 is a new prediction model constructed by us, which adopts a new mode of TCM knowledge question answering based on multi-hop reasoning and convolutional neural network combined with question answering based on TCM knowledge graph. It has four modules: (1) Problem embedding module; (2) Knowledge graph embedding module; (3) Link prediction scoring module, using convolutional neural network model InteractE link prediction; (4) Answer selection module combines multi-hop reasoning answer score and link score to select the most accurate answer for TCM knowledge problems. Fig.1 shows each module:

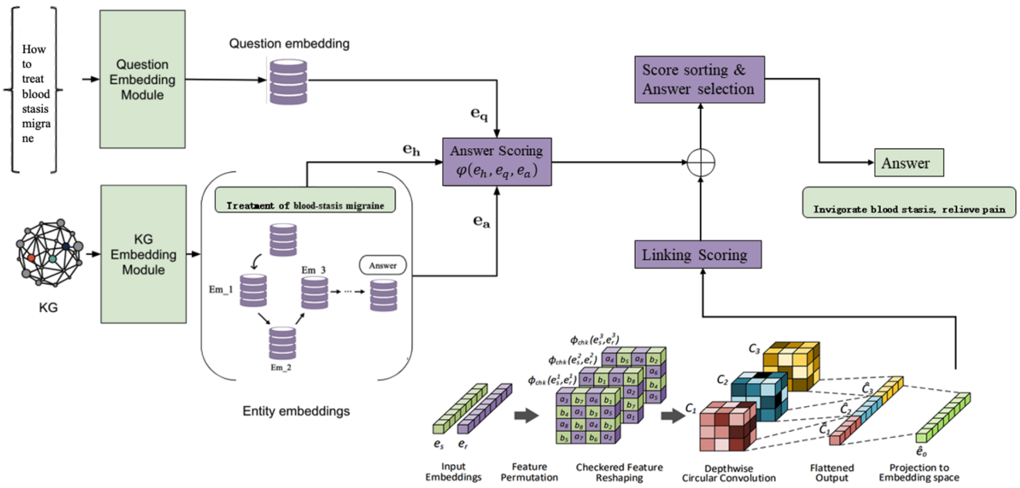

Fig. 1. Overview of EITCMQA model

2.2.1 Problem embedding module This module embeds the patient's consultation problem q into a fixed dimension vector $e_q \in C^d$. A feedforward neural network can accomplish this process. The first step is to embed the problem q into a 768-dimensional vector using Roberta.

We have a topic entity in TCM, a patient's self-report problem q, and learn the answer entity set $A \in \mathcal{E}$ embedded in the problem using formula (1) and formula (2).

$$\varphi(e_h, e_q, e_a) > 0 \quad \forall a \in A \tag{1}$$

$$\varphi(e_h, e_q, e_{\bar{a}}) < 0 \quad \forall \bar{a} \notin A \tag{2}$$

φ is described as equation (3), e_a, $e_{\bar{a}}$ are for the previous step of learning vector shallow.

$$\varphi(h, r, t) = R_e(\langle e_h, e_r, \bar{e_t} \rangle)$$
$$= R_e\left(\sum_{k=1}^{d} e_h^{(k)}, e_r^{(k)}, \bar{e}_t^{(k)}\right) \tag{3}$$

d stands for number of all relationships. For each patient's self-reported problem, the score φ (.) is calculated by all candidate answer entities $a_0 \in$ E. The model is learned by minimizing the binary cross entropy loss between the sigmoid score and the target tag, where the target tag is 1 or 0 of the correct answer.

2.2.2 Knowledge graph embedding module. The ComplEx embedding method involves training embeddings for all entites h, t $\in$ E, and relations r $\in$ R in TCM knowledge graph, such that $e_h, e_r, e_t \in \in C^d$. In the Q&A training set of TCM, according to the coverage of various entities in the knowledge graph of TCM, entities can be embedded and maintained and slightly changed.

2.2.3 InteractE convolution model. InteractE model is improved based on ConvE. The ConvE model first transforms the entity and relation vector representation into two dimensions, and performs convolution and other operations on them, and then obtains the corresponding probability through a full connection layer. InteractE model uses 2D convolution neural network (CNN) to model the interaction between entities and relationships. The score function used by ConvE is defined as:

$$\varphi(s, r, o) = f(vec(f([\bar{e}_s, e_r * w]))W)e_o \tag{4}$$

Among these, $\bar{e_s} \in \mathbb{R}^{d_w * d_h}$, $\bar{e_r} \in \mathbb{R}^{d_w * d_h}$ represents the 2D reconstruction of entity embedding $\bar{e_s} \in \mathbb{R}^{d_w d_h * 1}$ and (*) represents convolution operation. 2D reconstruction can enhance the interaction between entities and relational embedding, which has been proved to be helpful to learn better representation.

In order to capture various heterogeneous interactions, the model first generates t-random permutations of e_s and e_r, which are represented by $\mathcal{P}_t = [(e_s^1, e_r^1); \ldots \ldots; (e_s^t, e_r^t)]$. Note that for different i, the inner interaction set in $\Phi(e_s^i, e_r^i)$ is likely not to intersect.

Next apply remodeling operations $\Phi chk(e_s^i, e_r^i)$, $\forall i \in \{1, \dots, t\}$,and define $\Phi(\mathcal{P}_t) = [\Phi(e_s^1, e_r^1); \dots; \Phi(e_s^t, e_r^t)]$. ConvE uses $\Phi stk(\cdot)$ as a remodeling function with limited interactive capture ability. This model selects $\Phi chk(\cdot)$ as the remodeling function in InteractE, which captures the maximum heterogeneous interaction between entities and relational features. The formula of convolution in this way is expressed as:

$$[I \star w]_{p,q} = \sum_{i=-\lfloor\frac{k}{2}\rfloor}^{\lfloor\frac{k}{2}\rfloor} \sum_{j=-\lfloor\frac{k}{2}\rfloor}^{\lfloor\frac{k}{2}\rfloor} I_{[p-i]_m,[q-j]_n} w_{i,j} \tag{5}$$

The score function used in InteractE is defined as formula (6) :

$$\varphi'(e_h, e_q, e_a) = g(vec(f(\phi(\mathcal{P}_k) \circledast w))W)e_o \tag{6}$$

$\circledast$ represents the circular convolution in the depth direction, *vec(·)* represents the vector cascade, e_o represents the embedding matrix of the object entity, and *W* can learn the weight matrix. Select the functions *f* and *g* as ReLU and sigmoid. Standard binary cross entropy loss and label smoothing are used during the training.

***2.2.4 Answer Selection Module.*:**

However, for patients with multiple prescriptions, Chinese herbal medicine and other answer entities, we introduce the same link query mechanism in answer screening, such as formula (7) :

$$e_{ans} = arg_{a' \in \mathcal{E}} \max \varphi(e_h, e_q, e_{a'}) + \beta * \varphi'(e_h, e_q, e_{a'}) \tag{7}$$

According to knowledge graph embedding score and linking score, the answer entities are scored, and the optimal answer entities are given after ranking, in which β is the hyperparameter.

2.3 Experiment

2.3.1 Question Answering dataset. We validated our model using the cMedQA2 medical diagnostic dataset, which contains over 100,000 consultations related to male health, internal medicine, obstetrics and gynecology, oncology, pediatrics, and surgery in Chinese text. The dataset has three subsets: training, development, and test sets, with a total of 108,000 questions and 203,569 answers. On average, questions have 49 words, while answers have 101 words. More information about the dataset is available in Table 2, and details of the question-and-answer format can be found in Table 3.

***2.3.2 The experimental process and evaluation method*.** In reference to [4], the EmbedKGQA method was proposed, while references [6] and [7] proposed the ConvE and InteractE methods, respectively. These models were tested on the cMedQA2 medical question and answer dataset. The experimental results were evaluated based on the percentage of correct answers. To test the link prediction ability of the QA method, the experiment created an incomplete database by randomly discarding 25% and 50% of triples and conducted multiple experiments on the incomplete database. This was done to assess the performance of the knowledge graph embedding model when dealing with sparse data.

Table 2

The specific number of cMedQA2

Data set	Question	Answer	Average number of words in questions	Average number of words in answers
Training set	100000	188490	48	101
Development set	4000	7527	49	101
Test set	4000	7552	49	100
Total	108000	203569	49	101

Table 3

Examples of cMedQA2 dataset description

Question	Answer
I have hypertension these two days when my son-in-law came to give me some Pilosula to drink; hello hypertension, can eat Pilosula?	Hypertensive patients can take Codonopsis. Dangshen can lower blood lipid and blood pressure and eliminate the garbage in the blood, so it has a sound preventive effect on patients with coronary heart disease and cardiovascular disease so that oral Dangshen can be away from the harm of three highs.
Eyes dry, prickly, have the impulse of tears, usually work often on the computer, sleep quality is not very good. Dreams are easy to wake up.	According to the Chinese medicine theory, eye dry performance, to pay attention to conjunctivitis and other reasons caused by the possibility, belongs to the category of liver and kidney deficiency syndrome. Can use pearl eye drops, oral Qiju Dihuang pill to treat, do not drink, do not eat spicy and other irritant food, if the treatment effect is not good, should go to the regular hospital ophthalmology diagnosis and treatment.

To better evaluate our innovative models, we use the most commonly used metrics to measure the performance of knowledge graph embedding models: MRR, MR, HITS @ n. The calculation method for MRR is as follows:

$$MRR = \frac{1}{|S|}\sum_{i=1}^{|S|}\frac{1}{rank_i} = \frac{1}{|S|}\left(\frac{1}{rank_1}+\frac{1}{rank_2}+\cdots+\frac{1}{rank_{|S|}}\right) \tag{8}$$

S is the triple set, $|S|$ is the number of triple sets, and $rank_i$ refers to the link prediction ranking of the ith triple. The bigger the indicator, the better. For example, for the triple group (blood stasis headache, treatment, promoting blood circulation to remove blood stasis to relieve pain), the link prediction results are as Table 4. The results and answer for the treatment of the symptoms of dizziness and distension may be shown in Table 5.

Table 4

Link prediction results

s	p	o	score	rank
Headache of blood stasis syndrome	Therapies	Promoting blood circulation to remove blood stasis, promoting the restoration of special sense or consciousness and odynolysis	0.985	1*
Headache of blood stasis syndrome	Therapies	Clearing away heat and dampness, stasising pain	0.887	2
Headache of blood stasis syndrome	Therapies	Clearing away heat and promoting di-uresis, activating meridians to stop pain	0.856	3

Headache of blood stasis syndrome	Therapies	Air, activating blood	0.653	4

Table 5

Multi-hop reasoning results

s	p	o	score	rank
Headache of blood stasis syndrome	Therapies	Promoting blood circulation to remove blood stasis, promoting the restoration of special sense or consciousness and odynolysis	0.993	1*
Migraine	Therapies	Regulate liver Qi and tonify Qi and blood	0.709	2
Dizziness and bulge	Disease	Headache of blood stasis syndrome	0.653	3
Dizziness and bulge	Disease	Migraine	0.430	4

The calculation method for MR and HITS @ n is as formula (9) and (10):

$$MR = \frac{1}{|S|}\sum_{i=1}^{|S|} rank_i = \frac{1}{|S|}\left(rank_1 + rank_2 + \cdots + rank_{|S|}\right) \tag{9}$$

$$\text{HITS@n} = \frac{1}{|S|}\sum_{i=1}^{|S|} \mathbb{I}(rank_i \le n) \tag{10}$$

In general, n takes equal to 1,3 or 10. MRR and HITS @ 10 are two important indicators, and the larger the indicator is, the better it is. MR is not considered a good indicator. The smaller the indicator is, the better it is.

2.3.3 Experimental results. The experimental setup for this study involved using Windows 10 as the operating system and PyTorch with CUDA 10.1 for GPU acceleration.

Table 6 shows the evaluation results of each index using formula (8) - (10). To simulate performance in sparse databases missing triple information, we randomly discarded 25% and 50% triples to create an incomplete knowledge graph for TCM. The ComplEx, ConvE, and InteractE models were tested on this graph, with InteractE outperforming ComplEx and ConvE in most indicators.

Table 6

Experiment results

model	TCMKG				TCMKG(75%)				TCMKG(50%)			
	MRR	MR	HITS@10	HITS@1	MRR	MR	HITS@10	HITS@1	MRR	MR	HITS@10	HITS@1
ComplEx	.839	28.756	.910	.884	.796	60.327	.880	.758	.772	65.951	.828	.744
ConvE	.854	24.367	.916	.884	.808	58.829	.882	.765	.793	60.265	.834	.746
InteractE	.886	17.964	.935	.902	.812	54.489	.894	.783	.827	53.637	.836	.762

After evaluating the prediction results on the knowledge graph, we applied our model to a question-and-answer dataset that combines traditional Chinese and Western medicine. We combined our knowledge graph with an open-source medical knowledge graph from the Chinese Academy of TCM to make the results more accurate. To simulate a sparse graph, we randomly discarded 25% and 50% of the knowledge graph on cMedQA2. Our innovative model showed improved accuracy of about 2% compared to other models, and the accuracy of questions and answers is shown in Table 7.

Table 7

QA accuracy (Hits@1)

model	TCMKG	TCMKG(75%)	TCMKG(50%)
	HITS@1	HITS@1	HITS@1
EmbedKGQA + ComplEx	94.56%	78.40%	58.70%
EmbedKGQA +ConvE	95.98%	77.56%	59.08%
EmbedKGQA+ InteractE	96.55%	79.22%	61.59%

3. Conclusions

This paper introduces the EITCMQA method for TCM question answering, which combines convolutional neural networks and multi-hop reasoning with knowledge graph embedding. By increasing the dimension of the score function, this method provides a more accurate and comprehensive query path, effectively addressing the issue of missing answers in current knowledge graph embedding-based question answering methods. EITCMQA performs well on cMedQA2, particularly for multi-hop QA tasks and sparse knowledge graphs, outperforming baseline methods. Overall, our proposed method has clear advantages compared to existing methods.

Acknowledgments

This work is partially supported by Beijing Forestry University Science and Technology Innovation Program Project (BLX2014-27) , the National College Students' Training Programs of Innovation and Entrepreneurship (202110022062), and the CACMS Innovation Fund (CI2021A00512&CI2021A05403) .

REFERENCES

[1]. Yu Tong, Jia Lirong, Liu Jing, etc. Chinese medicine language system research review.Chinese Journal of Library and Information Science of Traditional Chinese Medicine, 2015,39(6):56-60.

[2]. Cao Mingyu, Li Qingqing, Yang Zhihao, etc.Knowledge Question Answering System for Primary Liver Cancer Based on Knowledge Map.Journal of Chinese Information, 2019, 33(6): 88-93.

[3]. Yang Yanling, Li Yan, Zhong Xinyu. A link prediction model for knowledge mapping of TCM medical records. TCM information,2022,39(03):1-6+15.DOI:10.19656

[4]. Odmaa, YANG Yunfei, SUI Zhifang, DAI Damai, CHANG Baobao, LI Sujian, ZAN Hongying. Preliminary Study on the Construction of Chinese Medical Knowledge Graph., 2019, 33(10): 1-7.

[5]. Zhou Hai, Peng Feiling, Wei Changfa. Research and practice of TCM diagnostic knowledge map construction.Medical Informatics Journal, 2020, 41(12):41-44.
[6]. Yin Dan, Zhou Lu, Zhou Yumei, etc. Research on the design of ' graph search mode ' of knowledge map of TCM classical prescriptions.Chinese Journal of TCM Information, 2019, 26(8):94-98.
[7]. Zhang Yuqi, Li Zongyou, Wang Yinghui, etc. Zhao Bingnan, Zhu Renkang Dermatology Schools with Fang Experience Knowledge Map Construction. Chinese Journal of Library and Information of Traditional Chinese Medicine, 2021, 45(2):1-5.
[8]. Liu Fan, Wang Mingqiang, Li Lingxiang, He Liyun. Exploration of the construction method of clinical experience knowledge map of famous and old Chinese medicine. Chinese Journal of Traditional Chinese Medicine, 2021,36(04):2281-2285.
[9]. Bordes A, Chopra S, Weston J. Question answering with subgraph embeddings. arXiv preprint arXiv:1406.3676, 2014.
[10]. Saxena A, Tripathi A, Talukdar P. Improving multi-hop question answering over knowledge graphs using knowledge base embeddings//Proceedings of the 58th annual meeting of the association for computational linguistics. 2020: 4498-4507.
[11]. Trouillon T, Welbl J, Riedel S, et al. Complex embeddings for simple link prediction//International conference on machine learning. PMLR, 2016: 2071-2080.
[12]. Dettmers T, Minervini P, Stenetorp P, et al. Convolutional 2d knowledge graph embeddings//Proceedings of the AAAI Conference on Artificial Intelligence. 2018, 32(1).
[13]. Shang C, Tang Y, Huang J, et al. End-to-end structure-aware convolutional networks for knowledge base completion//Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(01): 3060-3067.
[14]. Sun Z, Deng Z H, Nie J Y, et al. Rotate: Knowledge graph embedding by relational rotation in ComplEx space. arXiv preprint arXiv:1902.10197, 2019.
[15]. Vashishth S, Sanyal S, Nitin V, et al. Interacte: Improving convolution-based knowledge graph embeddings by increasing feature interactions//Proceedings of the AAAI conference on artificial intelligence. 2020, 34(03): 3009-3016.
[16]. Shi J, Cao S, Hou L, et al. TransferNet: An effective and transparent framework for multi-hop question answering over relation graph. arXiv preprint arXiv:2104.07302, 2021.
[17]. Liu Y, Hildebrandt M, Joblin M, et al. Neural multi-hop reasoning with logical rules on biomedical knowledge graphs//European Semantic Web Conference. Springer, Cham, 2021: 375-391.
[18]. Sun Mingjun, Zhang Dan, Zheng Mingzhi, etc. TCM auxiliary diagnosis and treatment system for rheumatoid arthritis based on artificial intelligence. Pattern recognition and artificial intelligence, 2021, 34(4): 343-352.
[19]. Liu, Zhang Yaquan. Assistant diagnosis of complications based on medical knowledge map. Journal of Chinese Information, 2020, 34(10): 85-93,104.
[20]. Gu Tianyu, Yan Zhuangzhi, Jiang Jiehui. TCM syndrome classification of stroke based on support vector machine. TCM information.