Intelligent Question Answering System Based on Domain Knowledge Graph

Yiming Hao[1a], Ye Wu[1b]*, Luo Chen[1, 2c] and Kaijun Yang[3d]

[1]*College of Electronic Science and Technology, National University of Defense Technology, Changsha, Hunan, China*
[2] *Key Laboratory of Natural Resources Monitoring and Supervision in Southern Hilly Region, Ministry of Natural Resources, Changsha, Hunan, China*
[3] *The Second Surveying and Mapping Institute of Hunan, Changsha, Hunan, China*
[a]*haoym_1998@163.com,* [b]*yewugfkd@nudt.edu.cn,* [c]*luochen@nudt.edu.cn,* [d]*apple_hacker@163.com*

Absrtact

This paper introduces an intelligent question answering system based on the domain knowledge graph of military battle cases. Through the collection and accumulation of military big data, we first build a domain knowledge graph for military battle cases, and then use natural language processing related technologies to understand natural language problems, mainly intention recognition and slot filling. On problem intent identification. In this paper, BERT+TextCNN model is proposed to realize the intention classification of questions. LAC is used to segment the natural language questions and extract the entities in the question sentence in slot filling. The answer is then retrieved from the knowledge graph. The test results show that the accuracy of question comprehension in the question set is more than 90%, and it can answer most of the questions in the field quickly and accurately.

Keywords: *Knowledge graph; Intelligent question answering system; Deep learning; Text classification*

1 INTRODUCTION

In 2012, Google put forward the concept of knowledge graph to improve its search quality [13]. Knowledge graph can conduct modeling and analysis for the objective world, and build a huge structured knowledge network of the entities of the objective world and the relationships among them [3].

The knowledge graph is divided into general knowledge graph and domain knowledge graph. The latter generally only focuses on the specific knowledge in a certain field, has obvious knowledge characteristics in the field and is easy to maintain.

With the increasing abundance of resources in the field of military cases, knowledge graph has gradually played a greater role in the field of military resources [5]. It is necessary to use multi-source knowledge to construct the domain knowledge graph of military cases [2]. Question-answering system based on knowledge graph has gradually become a hot research issue. Knowledge graph is an important driving force to transform data into knowledge and knowledge into action [7]. Research on intelligent question answering system based on knowledge graph will greatly improve the convenience of knowledge acquisition, and transform information advantage into knowledge advantage, and knowledge advantage into decision advantage [6].

2 RESEARCH STATUS

The earliest question-and-answer systems date back to the 1960s and 1970s. The 1961 BASEBALL system answered questions about a year's worth of BASEBALL games. The 1972 LUNAR system provided an interface for analyzing rock sample data during the Apollo moon missions. The early NLIDB emerged gradually with the development of artificial intelligence, transforming natural language problems into queries of structured knowledge base and obtaining answers. Its development benefited from many other development fields, involving expert system, linguistics, semantic web [10], database, information extraction and other fields, and was the product of cross fusion [9]. Then, the question answering system based on information retrieval (IRQA) is developed. According to the input questions, relevant documents are screened and key information is extracted as candidate answers by combining relevant technologies. Then, the candidate answers are sorted and the optimal

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 137-142, 2023.
https://doi.org/10.2991/978-94-6463-040-4_21

answers are returned. In the 1970s and 1980s, the question answering system based on knowledge base came into being, which transformed questions into structured query statements, inquired knowledge base and returned answers. Computation of semantic similarity between problems [12] belongs to short-text similarity calculation, which has always been a research hotspot in the field of NLP. Most of the key information exists in unstructured information. In order to improve the accuracy of question answering, unstructured information and structured information are often combined to form a hybrid question answering framework.

3 SYSTEM DESIGN

3.1 *System Design Process*

User input questions, the system analyzes the questions, mainly for slot filling and intention identification, and converts the results into structured query statements, query answers.

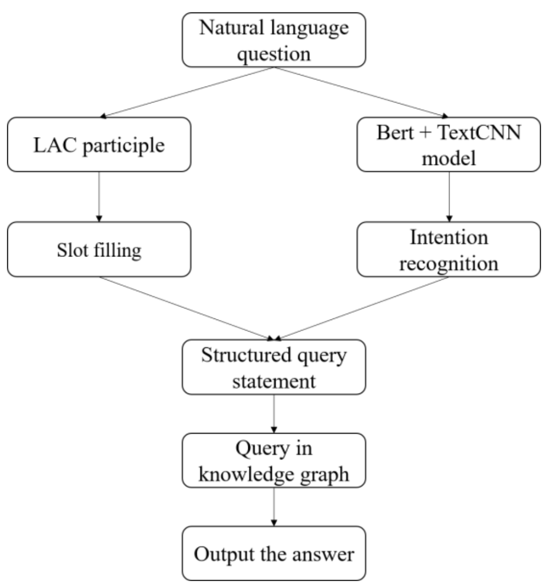

Figure 1 Flow chart of intelligent question answering system

3.2 *Implementation of Key Technologies*

3.2.1 Data Source of Knowledge Graph

The data used to construct the knowledge graph comes from network encyclopedia data, OpenKG, etc., and the multi-source data are integrated to construct the knowledge graph in the field of military battle cases. It contains battle cases, military figures, organizations, locations and other entity data. There are rich relationships between entities, and each type of entity contains rich attribute information.

3.2.2 Slot Filling

The task of slot filling is to identify the named entity in the user's question, find the subject that the user wants to query, and then query the relevant relationship or attribute according to the subject. This step mainly uses the LAC toolkit. LAC is a deep learning Chinese lexical analysis tool, which is a combined lexical analysis model. It can complete the tasks of Chinese word segmentation, part of speech tagging and proper name recognition. It is based on a stacked bi-directional GRU structure.

3.2.3 Intent Identification

After obtaining the slot entity of the question, the computer is also required to understand the intention of the question, that is, the intention recognition of the question is essentially a multi-classification problem of the text, which belongs to the knowledge category of natural language understanding. The quality of classification results is the key factor that restricts the retrieval effect of question answering system [8].

Firstly, we need to specify the common problem intentions in the field, secondly, we construct the question training data set, and train the data model, the common way is to use machine learning or deep learning methods, such as SVM algorithm or neural network model.TF-IDF or Word2vec model is usually used for text vector modeling, which can generally achieve better results, but it can not make full use of the meaning of words in the text. Therefore, this paper uses Bert model as the vector feature extractor of the training set text, maps the natural language questions into vectors rich in semantic feature information, and classifies the questions through TextCNN classification model. Matches to the specified intent library.

3.2.4 BERT+TextCNN Model

In the process of intention recognition, it is necessary to encode the vector features of the text. Because the BERT pre-training language model has strong semantic expression ability, it can accurately capture the semantic feature information from the sentence [1]. Therefore, this paper uses the bidirectional Transformer coding structure embedded in the model. The BERT pre-training language model structure is shown in Figure 2.

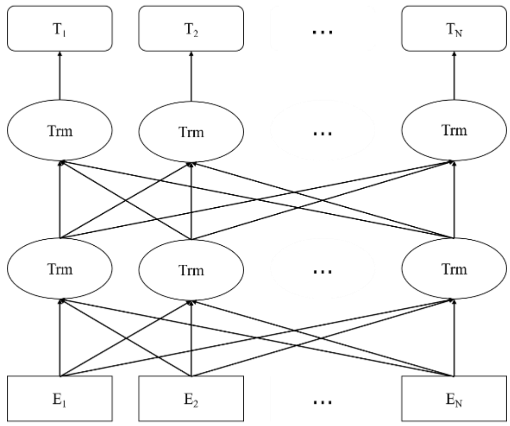

Figure 2 BERT pre-trained language model structure

BERT is a pre-training model, which learns a large number of prior sentences, syntax, word meaning and other information for downstream tasks through unsupervised training of a large number of corpus in the early stage. There is no need to decode to complete the specific task, so the BERT model mainly uses the Encoder in Transformer as the encoder, and does not use its Decoder for decoding. And BERT uses a bidirectional language model. It can better integrate the knowledge of the context.

The most important part of BERT is the bidirectional Transformer coding structure. Transformer abandons neural network structures such as CNN and RNN, and is an encoder-decoder model that completely uses self-attention mechanism to calculate input and output. It can process text in parallel [4] with high computational efficiency. The structure of the encoder is shown in Figure 3.

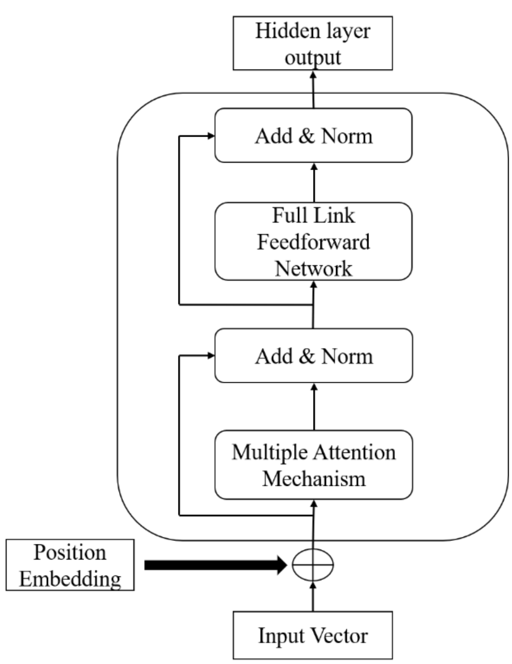

Figure 3 Transformer Encoding Unit

The coding unit is mainly divided into four parts: word vector and position coding, self-attention layer, residual connection and layer normalization, and fully connected feedforward network layer.

Word vector and position embedding: Transformer cannot recognize natural language, so it is necessary to convert language into word vector, that is, to convert text into mathematical representation recognized by the computer. Word2Vec can be used to convert text-> word vector. There is no cycle structure similar to RNN in Transformer, which can not capture the timing information of sequences, so it is necessary to introduce position coding information.

Self-Attention layer: With the word vector matrix and position embedding information, the most important Self-Attention layer in the coding unit is calculated, which mainly adjusts the weight coefficient matrix through the degree of association between words in the same sentence to obtain the representation of words. As shown in formula (1).

$$Attention\ (Q,K,V) = Softmax\ (\frac{QK^T}{\sqrt{d_k}})\ V \tag{1}$$

Where Q, K, V are all input word vector matrices and d_k is the input vector dimension.

Multiple attention mechanisms are shown in Equations (2) and (3).

$$MultiHead\ (Q,K,V) = Concat(head_1,\ head_2 \ldots,\ head_n)W^O \tag{2}$$

$$head_i = Attention\ (QW_i^Q,\ KW_i^K,\ VW_i^V) \tag{3}$$

The three weight matrices are divided, and the matrix obtained by calculating the dot product of the divided weight matrices is the attention matrix. Formula (1) is the self-attention mechanism. Each row of the matrix V represents the mathematical expression of each word vector. First, the attention matrix QK^T is obtained. Then the attention matrix is used to weight V.$\sqrt{d_k}$In order to change the attention matrix into a standard normal distribution, Softmax normalizes the sum of the attention weights of each word and all other words to 1.

Residual connection and layer normalization: After the operation of each module, the value before the operation is added to the value after the operation to obtain the residual connection. The main function of layer normalization is to normalize the hidden layer in the neural network to the standard normal distribution, which is convenient for speeding up the training speed. Accelerate convergence.

Fully connected feed-forward network layer: The feed-forward network layer is relatively simple and is a two-layer fully connected network. The activation function of the first layer is ReLU, and the activation function of the second layer is not used. The corresponding formula is shown in formula (4).

$$FFN(X) = \max(0, XW_1 + b_1)\, W_2 + b2 \qquad (4)$$

Where X is the input of the feedforward layer, and the dimension of the output matrix of this layer is consistent with X.

TextCNN is an application of convolutional neural network in text classification. Its core idea is to capture local features.The advantage of convolutional neural network is that it can automatically combine and filter N-gram features to obtain different levels of semantic information. TextCNN has four layers, and its structure diagram is shown in Figure 4.

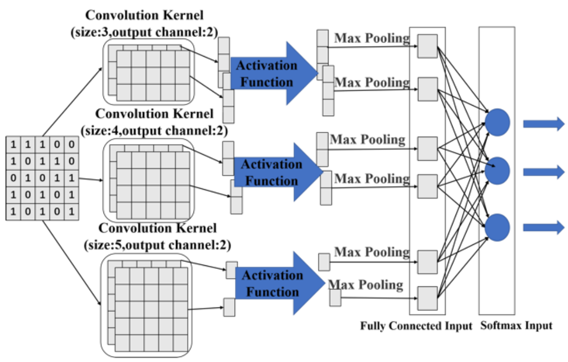

Figure 4 TextCNN structure diagram

The first layer is the input word vector layer, which takes the semantic word vector of the output of the BERT pre-training model as the input of this layer.

The second layer is the convolution layer. Since the text expressed by the word vector is one-dimensional data, the TextCNN convolution uses one-dimensional convolution, which acts on the word vector layer through multiple filters, and different filters generate different feature maps.The third layer is the maximum pooling layer .The fourth layer is the fully connected softmax layer, which splices the results of the maximum pooling layer and performs n classification through softmax .

The BERT+TextCNN deep learning model can accurately understand the intention of the question, and intention recognition is the core part of intelligent question answering. Only by correctly and efficiently understanding the key information such as the user's intention to ask a natural language question, can the follow-up knowledge retrieval be effectively carried out.

3.2.5 Question Conversion and Query

After getting the slot entity and the question intention by processing the user's question, the next step is to convert it into a query statement that can be recognized by the knowledge base. This enables the leap from unstructured natural language to structured query statements [11]. The system uses Neo4j graph database to store knowledge graph. When the user enters a question, the system will analyze the question sentence. And extract that slot position and the identification intention, mapping the slot position and the identification intention into a query statement to perform result query, and perform answer feedback.

4 EXPERIMENT

Next, the performance of BERT+TextCNN deep learning model in natural language problem intention recognition task is verified through experiments, and its performance is compared with other common intention recognition models.

4.1 Dataset

The data set selected in this experiment is the short text data of Toutiao official website. It covers 15 categories, including cultural field, entertainment field, sports field and economic field. 382657 sample data were obtained through data preprocessing. The numb of samples for each category and that correspond category labels are shown in Table 1. In order to verify the accuracy and validity of the model, the data set was randomly divided, 90% of the data was selected as the training set, and the remaining 10% was used as the test set.

Table 1: Sample Category and Quantity Statistics

Categories	Quantities
news_story	6273
news_culture	28031
news_entertainment	39395
news_sports	37568
news_finance	27085
news_house	17672
news_car	35785
news_edu	27030
news_tech	41542
news_military	24984
news_travel	21422
news_world	26908
news_stock	340
news_agriculture	19322
news_game	29300

4.2 Experimental Comparison

In this experiment, three deep learning models are constructed to recognize the intention of natural language problems, and their performance is compared. They are TextRNN model, TextRCNN model and BERT+TextCNN model. The input layer word vectors of the TextRNN model and the TextRCNN model are generated by word2vec. Table 2 shows the comparison of classification effects of different intention recognition models.

Table 2 Performance Comparison of Different Intention Recognition Algorithms

Algorithm	precision	recall	f1-score
BERT+TextCNN	0.90	0.88	0.89
TextRNN	0.84	0.82	0.83
TextRCNN	0.68	0.70	0.69

5 CONCLUSION

In view of the current research status in the field of intelligent question answering, the method of deep learning is introduced into the question answering system, and an intelligent question answering system for military battle cases based on domain knowledge graph is realized. Through deep learning method, select the appropriate training set, train the corresponding characteristics can accurately understand the user's intention, to achieve intelligent question and answer.

6 REFERENCES

[1] Chen Ping et al. Intelligent Question Answering System by Deep Convolutional Neural Network in Finance and Economics Teaching[J]. Computational Intelligence and Neuroscience, 2022, 2022: 5755327-5755327.

[2] DOU Xiao-qiang and LIU Tian-ya and ZHANG Zhi-zheng. Question Answering System Based on Military Knowledge Graph [C]. Proceedings of the 6th China Command and Control Conference (Volume 1), 2018:537-541.

[3] Fu Shan and Lu Ailin and Yan Shu. The Concept and Application of Knowledge Graph [J]. Information and Communications Technology and Policy, 2019(05):10-13.

[4] Gan Ling and Xiao Yang. Knowledge Base Question Answering Based on Multi-head Attention Mechanism and Relative Position Coding [J]. Journal of Physics: Conference Series, 2022, 2203(1)

[5] GE Bin and TAN Zhen and ZHANG Chong and XIAO Wei-Dong. Military Knowledge Graph Construction Technology [J]. JOURNAL OF COMMAND AND CONTROL, 2016,2(4):302-308.

[6] Jiang Kai and Qian Kui and Zheng Xuan. Military Information Retrieval Technology Architecture Based on Knowledge Graph [J]. Command Information System and Technology, 2016, 7(01): 47-52. DOI: 10.15908/ j. cnki. cist.2016.01.010.

[7] LAI Rong-xuan and ZOU Shun and WU Wen-hui and WU Ze-xin. Review of the Application of Knowledge Graph in the Field of Military Human Resources [J]. Software Guide, 2021,20(06):243-247.

[8] Limin Shao and Qingxia Xue and Xiaofang Liu. Question Classification of Question Answering System in Financial Management Based on BERT [J]. Scientific Journal of Intelligent Systems Research, 2021, 3(10)

[9] Li Tao and Wang Cichen and Li Huakang. Development and construction of knowledge graph [J]. Journal of Nanjing University of Science and Technology, 2017,41(01):22-34.

[10] Sowa J F.Principles of Semantic Networks: Exploration in the Representation of Knowledge [J]. Frame Problem in Artificial Intelligence, 1991(2-3): 135-157.

[11] Sui Yuan. Question answering system based on tourism knowledge graph[J]. Journal of Physics: Conference Series, 2021, 1883(1)

[12] Viji, D. and Revathy, S.. A hybrid approach of Weighted Fine-Tuned BERT extraction with deep Siamese Bi – LSTM model for semantic text similarity identification[J]. Multimedia Tools and Applications, 2022, :1-27.

[13] ZHU Muyijie, et al. Research progress on development and construction of knowledge graph [J]. Journal of Nanjing University of Information Science and Technology (Natural Science Edition), 2017, 9 (6):575-582.