Construction and Implementation of Tibetan Cultural Relics Question and Answer System Based on Knowledge Graph

Qingbei Chen [1,a], Liu Wang [1], Haiqiong Liang [2], Yiwei Zheng [3], Siyu Sun [4], Jiale Zhao [5], and Chang Liu [1]

[1] College of Information Science and Technology, Tibet University, Lhasa, Tibet 850000;

[2] College of engineering, Tibet University, Lhasa, Tibet 850000；

[3] College of science, Tibet University, Lhasa, Tibet 850000；

[4] College of Economics and Management, Tibet University, Lhasa, Tibet 850000；

[5] College of finance and economics, Tibet University, Lhasa, Tibet 850000；

[a]1357127082@qq.com,

Abstract. Tibet is a region with a long history and culture, with a large number of precious cultural relics resources. In the digital age, building a question-and-answer system based on the Knowledge map of Tibetan cultural relics can better protect and pass on Tibetan cultural heritage. Firstly, the Tibetan cultural relics data were obtained by crawler technology and stored in neo4j database to complete the knowledge graph construction. Secondly, in order to make the system understand the semantic information of the user's question, the bert_bilstm_crf model is used to recognize the entity information of the question, and the bert_textcnn model is used to recognize the intention of the question, and finally the answer result is returned by constructing the cypher query graph.

Keywords: Knowledge graph;Question answering;system;bert_bilstm_crf;bert_textcnn.

1. Introduction

Nowadays, the data on the Internet is growing at an exponential rate. When we browse the web, we often get lost in the unknown data, which is a puzzling thing for those who want to know the history, culture and cultural relics of Tibet. In view of this situation, this paper will study the Tibetan cultural relics question and answer system, so that more people can understand the Tibetan cultural relics in the form of question and answer.

2. The construction of question answering system

The question answering system based on knowledge graph constructed in this paper is divided into three processes: 1 knowledge graph construction, 2 question semantic parsing and 3 answer generation[1]. After the user enters the question, the system analyzes the entity information and intention information of the question, and queries the cultural relic atlas by constructing cypher query statement to return the question and answer result. In the following, the specific implementation of the three modules will be described respectively.

2.1 Construction of knowledge map of Tibetan cultural relics

A Knowledge Graph is a structured representation of knowledge that represents entities, attributes, and the relationships[2],it can be divided into two categories, vertical and general.General knowledge graph emphasizes the breadth of knowledge, such as Dbpedia[3], FreeBase[4], etc., while Vertical graphs emphasize depth of knowledge. For example, the knowledge graph of archival domain constructed by Wang Dianhua et al. [5] and the small financial knowledge graph constructed by Zhang Delang [6], etc., Building a knowledge graph is divided into several aspects , including knowledge extraction, knowledge fusion, knowledge processing and knowledge storage of data from different sources, and finally the knowledge graph is formed.

The data in this paper comes from the data of Tibet cultural relics on the Crawl network, including eight categories: steles, sculptures, calligraphy, paintings, ceramics, bronze ware, embroidery, and religious cultural relics. After data cleaning, a total of 8067 structured data were obtained. The data was stored in csv file format, with the name of the cultural relic as the main body, and the relationship was established with the introduction of the cultural relic, the origin time of the cultural relic, the value of the cultural relic, and the address of the cultural relic unearthed. The data was stored in the neo4j database, Finally, the construction of knowledge graph is completed.

2.2 Semantic parsing of questions

Natural language understanding is one of the core tasks in the field of artificial intelligence, which aims to enable computers to understand and process human natural language. In the cultural relics question answering system, the semantic analysis of questions is divided into two sub-tasks: intention recognition and entity recognition. Entity recognition is to obtain the entity information of questions, and intention recognition is to determine the user's intention information.

2.2.1 named entity identification

In this paper, bert_bilstm_crf model is used to identify the names of cultural relics. The model as shown in Figure 1:

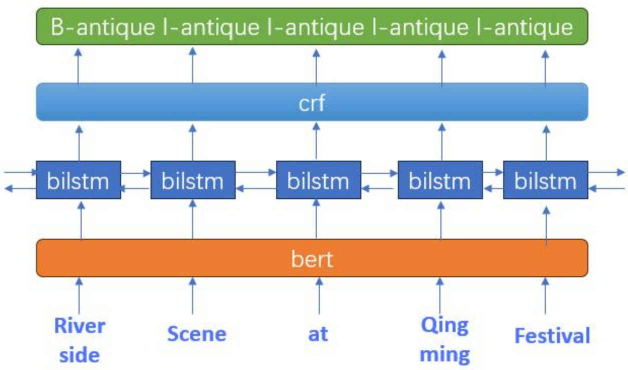

Figure1 bert_bilstm_crf model

Firstly, the pre-trained BERT[7] model is used to encode the input text, thus capturing rich semantic information. Next, the encoded text is modeled using bidirectional LSTM[8] to capture the sequential relationship of words in context. Finally, a conditional random field (CRF) [9] layer is used to transform the output of BiLSTM into a final named entity tag sequence, taking into account the dependencies between the tags. The model can accurately identify name entities such as personal name, place name and organization name in different contexts, and has high performance and generalization ability

BERT is a natural language processing model using a Transformer architecture with a bidirectional encoder that allows the model to consider both left and right lexical information in context. BERT uses large amounts of text data in the pre-training phase to learn a generic language representation, and then fine-tunes it on a specific task through fine tuning. Different from the traditional one-way Language Model, BERT used the Masked Language Model task to predict the masked words and the Next Sentence Prediction task to judge whether two sentences are adjacent. This pre-training allows BERT to capture the rich relationships between words and sentences

BiLSTM is a neural network architecture for sequential data modeling. It combines the idea of LSTM and bidirectional processing to capture temporal dependencies in sequences. BiLSTM processes sequence data by encoding the input sequence in both forward and backward directions, with an LSTM layer in each direction. This allows the network to consider both the preceding and the following context information when predicting the output of a location, thus better extracting features, and BiLSTM preserves hidden states in both the forward and backward directions at each time step, which contain information about the past and future. By fusing this information, BiLSTM is able to capture more complex patterns in sequence data, improving the model's performance when processing sequence tasks.

CRF is designed to solve the label dependency modeling problem in sequence data. In many sequence labeling tasks, there is a dependency between labels, that is, the prediction of the current label may be influenced by the previous label. CRF better captures these dependencies by considering the probability of transitions between global features and labels. In CRF, the input features and output labels of a sequence are modeled as a conditional probability distribution. The goal of the model is to maximize the probability of the output sequence given the input conditions. The core idea of CRF is to define two kinds of probabilities: observation probability and transfer probability. The observed probability represents the probability of a label at each location given an input. The transition probability represents the conversion probability between labels and is usually learned from the model parameters. The CRF model adjusts these parameters during training to maximize the probability of observing the corresponding label sequence.

2.2.2 Text classification

In this paper, the Bert_TextCNN model is used to classify user intentions. TextCNN[10] is a deep learning model widely used in the field of text classification. By drawing on the idea of CNN, it can effectively extract features from text data to achieve accurate text classification. The core structure of TextCNN includes convolution layer and pooling layer. The model first converts the words in the text into corresponding word vectors, and then uses one-dimensional convolution kernel to slide over these word vectors to capture local features of different sizes. The convolution operation can effectively capture the relationship between words and help the model understand the semantic structure of the sentence. Next, the most significant portion of the features generated by each convolution kernel is selected through a maximum pooling operation to reduce the dimension of the features while preserving important information. Finally, the pooled features are connected to the fully connected layer for the final classification

2.3 Answer generation

Build a query template， match （n： antique） where n.antique= ‘{name}’ return n.‘{intention}’， By replacing the query template‘name’、‘intention’，construct the final query statement, and return the question and answer result through the query graph。

3. Experiment

3.1 Experimental parameter settings

In bert_bilstm_crf model, max_len is set to 150, batch_size is set to 32, lstm_units is set to 128. In order to prevent overfitting, drop_rate is set to 0.1, and learning rate is set for layers. bert layer does not participate in training. Set the other layers to 0.0001 and train 50 epochs.

In the bert_Textcnn text classification model, class_num is set to 4, maxlen is set to 128 and batch_size is set to 32. The convolutional layer extracts local features, and the pooling layer gets the maximum feature value. After three convolutional pooling operations, the three maximum values are spliced into the fully connected layer for prediction.

3.2 Analysis of results

In order to compare the performance of various models in the recognition of cultural relics named entities and text classification, this paper uses precision rate, recall rate and F1 value to measure the performance of different models.

The bert_bilstm_crf model is compared with lstm, lstm_crf and bilstm_crf, and the experimental results are shown in Table 1:

Table 1 comparison of models

Models	P	R	F1
lstm	70.06%	72.14%	71.08%
lstm_crf	76.51%	77.32%	76.91%
bilstm_crf	80.26%	84.87%	82.50%
bert_bilstm_crf	89.48%	90.47%	89.97%

It can be seen from the table that the F1 value of the lstm_crf model is 5.83% higher than that of the lstm model, indicating that the dependency relationship between label sequences can be modeled after crf is added. In sequence labeling tasks, especially named entity recognition tasks, labels often have a certain context dependence. CRF can optimize the tag sequence on a global scale, making the conversion between adjacent tags more reasonable. F1 value of bilstm_crf model is 5.59% higher than that of lstm_crf model, indicating that BiLSTM (bidirectional short and long term memory network) can consider both forward and backward context information. Compared with traditional LSTM which only considers forward information, BiLSTM can capture the dependencies in sequence data more comprehensively. The F1 value of the bert_bilstm_crf model used in this paper is 7.47% higher than that of the bilstm_crf model, indicating that BERT model learns rich semantic context information through large-scale corpus in the pre-training stage. This allows the model to understand the semantics of the text more deeply and to better capture dependencies between contexts.

The bert_textcnn model used in this paper is compared with textcnn and bert_softmax, and the experimental results are shown in Table 2:

Table 2 Comparison of text classification models

Models	P	R	F1
textcnn	76.24%	70.48%	73.25%
bert	80.62%	78.57%	79.58%
bert_textcnn	88.86%	86.87%	87.85%

It can be seen from the table that the F1 value of the bert pre-trained model is 6.33% higher than that of the textcnn model, indicating that BERT can deeply understand the context of text through bidirectional modeling. In contrast, TextCNN mainly extracts local features through convolution operations and lacks global context understanding. The F1 value of the Bert_textcnn model used in this paper is 8.27% higher than that of the bert model, indicating that bert_textcnn combines the context understanding of Bert and the local feature capture capability of TextCNN, which helps to capture text information more comprehensively and improve classification performance.

4. Summary

In this paper, the knowledge map of Tibetan cultural relics is constructed first. Secondly, the entity information is identified by the bert_bilstm_crf model, and the intention information is identified by the bert_textcnn text classification model, so as to analyze the semantics of user questions. Thirdly, the two kinds of information are used to identify the intention of the query statement and fill the slot, and finally complete the construction of the question answering system.

References

[1] Xia Chengkui. Design and implementation of government Knowledge graph question answering system based on deep learning [D]. Guizhou University.2022.001005.

[2] Guo Lin, Chen Xiaohui, Xiao Mei. Information Recording Materials,2023,24(06):17-19+ 23.06.006. (in Chinese)

[3] Bizer C, Lehmann J, Kobilarov G, et al. DBpedia - A Crystallization Point for the Web of Data[J]. Web Semantics: Science, Services and Agents on the World Wide Web, 2009,7:154-165.

[4] Bollacker K, Evans C, Paritosh P, et al. Freebase: A collaboratively created graph database for structuring human knowledge[C]//Proceedings of the 2008 ACM SIGMOD international conference on Management of data, 2008:1247-1250.

[5] Wang Dianhua, Qian Tao, Qian Lixin, et al. Research on archive-oriented Knowledge graph construction method [J]. Journal of Hubei University of Science and Technology, 20, 40 (1) : 127-130. (in Chinese

[6] Zhang Deliang. Research on Construction and Completion of Knowledge Graph for financial field [D]. Qingdao: Qingdao University, 2020,10(3):36-48. (in Chinese

[7] BERT: Pre-training of Deep Bidirectional Transformers for Language nderstanding.2018

[8] Hochreiter,S,Schmidhuber,J.Long,short,term,memory[J].NeuralComputation,1997,9(8):1735-1780.

[9] Lafferty J.D, Mccallum A.K, Pereira F.C.N. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data[J]. 2001

[10] Yin W P,Schütze H,Xiang B,et al.ABCNN:attention-based con ▪ volutional neural network for modeling sentence pairs[J].Trans ▪ actions of the Association for Computational Linguistics,2016,4:259-272.