information

MDPI

Article

Preliminary Study on the Knowledge Graph Construction of Chinese Ancient History and Culture

Shuang Liu [1,*], Hui Yang [1], Jiayi Li [1] and Simon Kolmanič [2]

[1] School of Computer Science and Engineering, Dalian Minzu University, Dalian 116600, China; yanghuimail2020@gmail.com (H.Y.); ygrace788@gmail.com (J.L.)
[2] Faculty of Electrical Engineering and Computer Science, University of Maribor, Koroska cesta 46, SI-2000 Maribor, Slovenia; simon.kolmanic@gmail.com
* Correspondence: liushuang@dlnu.edu.cn; Tel.: +86-156-4119-6451

Received: 23 March 2020; Accepted: 28 March 2020; Published: 30 March 2020

Abstract: The domestic population has paid increasing attention to ancient Chinese history and culture with the continuous improvement of people's living standards, the rapid economic growth, and the rapid advancement of information science and technology. The use of information technology has been proven to promote the spread and development of historical culture, and it is becoming a necessary means to promote our traditional culture. This paper will build a knowledge graph of ancient Chinese history and culture in order to facilitate the public to more quickly and accurately understand the relevant knowledge of ancient Chinese history and culture. The construction process is as follows: firstly, use crawler technology to obtain text and table data related to ancient history and culture on Baidu Encyclopedia (similar to Wikipedia) and ancient Chinese history and culture related pages. Among them, the crawler technology crawls the semi-structured data in the information box (InfoBox) in the Baidu Encyclopedia to directly construct the triples required for the knowledge graph, crawls the introductory text information of the entries in Baidu Encyclopedia, and specialized historical and cultural websites (history Chunqiu.com, On History.com) to extract unstructured entities and relationships. Secondly, entity recognition and relationship extraction are performed on an unstructured text. The entity recognition part uses the Bidirectional Long Short-Term Memory-Convolutional Neural Networks-Conditions Random Field (BiLSTM-CNN-CRF) model for entity extraction. The relationship extraction between entities is performed by using the open source tool DeepKE (information extraction tool with language recognition ability developed by Zhejiang University) to extract the relationships between entities. After obtaining the entity and the relationship between the entities, supplement it with the triple data that were constructed from the semi-structured data in the existing knowledge base and Baidu Encyclopedia information box. Subsequently, the ontology construction and the quality evaluation of the entire constructed knowledge graph are performed to form the final knowledge graph of ancient Chinese history and culture.

Keywords: knowledge graph; ancient history and culture; knowledge extraction; named entity recognition; visual display

1. Introduction

As of June 2019, the number of Internet users in China has reached 854 million, an increase of 25.98 million compared with 2018, according to the 43rd Statistical Report on the Development of the Internet in China issued by the China Internet Network Information Center (CNNIC) [1]. As of June 2019, the total number of websites in China has reached 5.18 million. Among the many websites, the number of history and culture related websites in China has dramatically increased. In particular, websites that

are related to ancient Chinese history and culture have grown at an alarming rate. This means that more and more people have begun to pay attention to history and culture and understand the cultural heritage of our country for the past five thousand years.

Chinese ancient history is very important intangible cultural heritage in the long history of the Chinese nation's culture. With the rapid development of the Internet, there are massive multimedia historical and cultural data of ancient China in the format of video, audio, and texts on the Internet. How to utilize these data efficiently with scientific methods is our focus. Organizing and protecting ancient Chinese history is the inheritance and promotion of the 5000 years of Chinese culture.

In the era of big data, the knowledge graph is an important data resource for knowledge management and application. It has become a key technology in application fields, such as semantic retrieval, knowledge reasoning and decision making, knowledge answering, and recommendation systems. In this context, Google introduced the concept of Knowledge Graph (KG) in 2012, which aims to improve the performance of search engines [2]. The knowledge graph is stored in the form of structured triples. The basic unit of composition is composed of the head entity, the tail entity, and the relationship between the two entities. The general expression is $G = (E, R, S)$, where $E = \left\{e_1, e_2, e_3 \ldots, e_{|E|}\right\}$ represents the entity set, $R = \left\{r_1, r_2, r_3 \ldots, r_{|R|}\right\}$ represents relation set, and $S \subseteq E \times R \times E$ represents the set of triples in the knowledge graph.

Presently, the research and application of knowledge graphs are mainly divided into the general domain knowledge graph and vertical domain knowledge graph. The general domain knowledge graph in English includes YAGO [3], DBpedia [4], Wikidata [5], etc. A typical Chinese general domain knowledge graph includes CN-DBpedia [6], zhishi.me [7], Ownthink [8], XLore [9], etc. The knowledge graph of vertical industry domains in English includes IMDB [10], MusicBrainz [11], etc. The knowledge graph of Chinese vertical industry domains includes the Traditional Chinese Medicine (TCM) knowledge graph [12], marine knowledge graph, and corporate knowledge graph. Although the above general domain knowledge graph collects a large amount of domain knowledge, it cannot describe the knowledge in a certain field in detail. The advantages of the vertical domain knowledge graph in this respect are greater than the general domain knowledge graph, but the knowledge graph construction in this field is often manually constructed, which requires a considerable amount of human and financial resources.

After investigation, it is found that the existing knowledge graph of the general domain contains some contents that are related to Chinese ancient history and culture, but the existing knowledge graph has a lot of room for improvement in terms of scale, standardization, and formalization [13]. Presently, there is no knowledge graph of Chinese ancient historical and cultural knowledge in the vertical field. How to build a large-scale and high-quality graph of Chinese ancient historical and cultural knowledge based on efficient knowledge engineering methods and advanced text data mining technology is still a very challenging subject.

In view of the above challenges, this paper uses natural language processing (NLP), text data mining technology, and knowledge graph construction technology to build one Chinese ancient history and culture knowledge graph by using a combination of manual method and machine automation method.

The main purpose of this article is to implement a knowledge graph construction model to provide better knowledge services for applications that are related to ancient Chinese history and culture. The knowledge that is provided by our model can be used in many aspects, such as recommendation systems and question answering systems. In summary, the main contributions of this article are as follows:

- This paper proposes a model for constructing a graph of knowledge regarding ancient Chinese history and culture.
- It introduces, in detail, how the knowledge graph is constructed, and the main steps in the construction process, namely entity recognition.

- Finally, a visual display of the ancient Chinese history and culture knowledge graph constructed is convenient for the public to better understand ancient history and culture.

The rest of the paper is organized, as follows. Section 2 introduces the related work of knowledge graph construction and named entity recognition. Section 3 introduces the overall scheme of construction of our Chinese ancient history and culture knowledge graph. Section 4 introduces knowledge extraction technology, mainly entity extraction. Section 5 gives the experimental results and analysis. Section 6 is the visualization system display of ancient Chinese history and the culture knowledge graph. Section 7 is the summary of the paper and future outlook.

2. Related Work

2.1. Knowledge Graph Construction

Many researchers have done a lot of work to build large-scale, high-quality, general-purpose knowledge bases, such as the aforementioned YAGO, DBpedia, Wikidata, CN-DBpedia, zhishi.me, and Ownthink, since the concept of knowledge graph was proposed in 2012. However, the knowledge graph of the general domain is often difficult to cover the knowledge of the professional domain. The construction of vertical industry knowledge graph has become a research hotspot in order to make the knowledge graph more suitable for some professional fields.

In the field of ancient history and culture, research scholars have developed various historical and cultural-based ontology and knowledge bases to achieve knowledge management and knowledge sharing. Laura Pandolfo proposed in STOLE: A Reference Ontology for Historical Research Documents, a reference ontology to build a historical ontology library in Italian public administration [14]. Martin Doerr constructed a knowledge base in the field of cultural heritage that is based on ontology in Ontologies for Cultural Heritage [15]. Zhou et al. constructed a historical figure knowledge graph in a big data environment, and visualized the obtained data through visualization technology [16].

In other areas, the Gene Ontology [17] constructed Gene Ontology, which can be used to describe the genes and gene products in any organism. Qiu Minghu et al. introduced the construction of the recipe ontology in detail, mainly including the daily recipe and food composition database [18]. Ruan et al. established an open knowledge base of traditional Chinese medicine symptoms that can be used in clinical decision support systems, including concepts of diseases, drugs, and the relationship between symptoms and the above entities [19]. The Computer Knowledge Engineering Lab of Tsinghua University constructed a bilingual knowledge graph of film and television, which is mainly integrated with LinkedIMDB, Baidu Encyclopedia, Douban (a website that offers recommendations, reviews, and price comparisons for books, movies, music records, and the city's unique cultural life), and other data sources [20]. Vrije University of Amsterdam constructed a breast cancer knowledge graph, which mainly integrates breast cancer related knowledge [21]. The Chinese Academy of Chinese Medical Sciences constructed a medical knowledge graph, which mainly includes a knowledge graph of TCM medical records, and a knowledge graph of TCM characteristic diagnosis technology [12].

These constructed knowledge graphs have made great contributions to applied research in this direction. However, due to the complexity and variety of data in ancient Chinese history, it is different from data in other fields with relatively uniform formats. Data integration in the field of history and culture is relatively difficult. Accordingly, our research goal is to establish a large-scale, high-quality knowledge graph of ancient Chinese history and culture, laying the foundation for research in the field of history and culture.

Presently, although there are a lot of knowledge graphs in vertical fields, most of them are constructed by ontology fusion and database integration. The cultural heritage knowledge graph constructed by Martin Doerr is based on a relational database and a hierarchical database system. The data source of the STOLE knowledge graph is mainly historical texts and literature, including periodicals and newspapers at that time. Zhou Yi et al. constructed a knowledge graph of historical

figures, while using data from Google, Baidu, Sogou (a website), and other websites. The field of knowledge graph is constructed by it is relatively single, only historical figures.

First of all, this article is different from the data sources of the cultural heritage knowledge graph, historical knowledge graph (STOLE), and historical figures at public administration in Italy. The knowledge graphs that are mentioned above are all based on ontology fusion or database integration. The data source of the knowledge graph that was constructed in this paper is to extract and integrate knowledge from various Internet data sources (including structured data, semi-structured data, and unstructured data) [22]. Secondly, the content of these existing knowledge graphs that are related to history and culture is relatively simple. In contrast, the model can automatically extract knowledge from unstructured text information, which makes it possible to integrate knowledge from more open and diverse data sources [23]. In addition, from the perspective of knowledge, we have introduced some historical and cultural-related knowledge, including entities, such as dynasties, characters, and war, and the relationships between them into the ancient Chinese historical and cultural knowledge graph to make it more comprehensive. Finally, most of the current knowledge graphs in the field of history and culture are in English, which might not be suitable for Chinese reading habits. We have established a Chinese history and knowledge graph. This study constructs the ancient Chinese history and culture into a knowledge graph to help the public to better understand the ancient Chinese history and culture.

2.2. Named Entity Recognition

With the rapid development of Internet technology, people began to pay attention to entity extraction in the vertical domain while performing entity extraction on general domain data. However, text data in the vertical domain has its own characteristics, and its own characteristics need to be considered when performing entity extraction [24].

In the work of named entity recognition, it is mainly divided into rule-based methods, statistical machine learning-based methods, and neural network-based methods. Among the common statistical machine learning-based models are Hidden Markov Model (HMM) [25], Maximum Entropy Model [26], Maximum Support Vector Machine (SVM) [27], and Conditions Random Field (CRF), etc. [28]. However, these methods need to be manually done when performing feature extraction. At the same time, a lot of manual labeled samples are needed in model training, and the effect is not obvious [29].

The method that is based on neural network is usually regarded as a sequence labeling task in named entity recognition tasks, and the text is used for entity recognition by establishing a sequence labeling model. In 2011, Collobert et al. used Convolutional Neural Networks (CNN) for feature extraction, and achieved good recognition results by fusing other feature effects [30]. In 2015, Huang et al. proposed a Bidirectional Long Short-Term Memory-Conditional Random Field (BiLSTM-CRF) model to improve the model performance [31]. Santos et al. proposed using the character CNN to enhance the CNN-CRF model [32]. In 2016, Lampleet et al. used two BiLSTMs to learn word-level and character-level features, respectively [33]. In 2017, Strubell et al. proposed the use of void convolutional network (IDCNN-CRF) for named entity recognition to extract sequence information and accelerate the training speed [34]. In 2018, Fenget et al. proposed a named entity recognition method that was based on BILSTM neural network structure [35]. Maimatiayifu et al. proposed a BiLSTM-CNN-CRF model [36]. Li Lishuang et al. applied the CNN-BILSTM-CRF model to biomedical corpus and obtained the highest F1 value at that time, according to the characteristics of the Uyghur language [37].

In this paper, the BiLSTM-CNN-CRF deep neural network model will be used in Chinese ancient historical and cultural entity recognition. In this process, Continuous Bag-of-Words Model (CBOW) will be used to train word vectors, convolution neural network will be used to extract the character representation vectors in sentences, character representation vectors, and word vectors will be spliced, and the spliced results will be used as input into Bidirectional Long Short-Term Memory (BiLSTM). Finally, the best annotation sequence will be selected by CRF according to the characteristics of the text to obtain the last identified entity information.

3. Knowledge Graph Construction Process

Many researchers have divided the construction process into several parts in the process of constructing the knowledge graph. Yang Siluo et al. divided the knowledge graph construction process into eight parts, which are sample data collection, sample data cleaning, knowledge unit selection, unit relationship construction, data standardization, sample data simplification, knowledge visualization, and results interpretation [38]. Katy Borner et al. divided it into six steps: extract data, define analysis units, select methods, calculate similarity, build knowledge units, and analyze results [39]. Although the process of constructing the knowledge graph is slightly different, they all mention the most important parts in the construction of the knowledge graph: data acquisition, information extraction, knowledge fusion, and graph construction. Figure 1 shows a flowchart of the construction of the graph of our Chinese ancient history and culture.

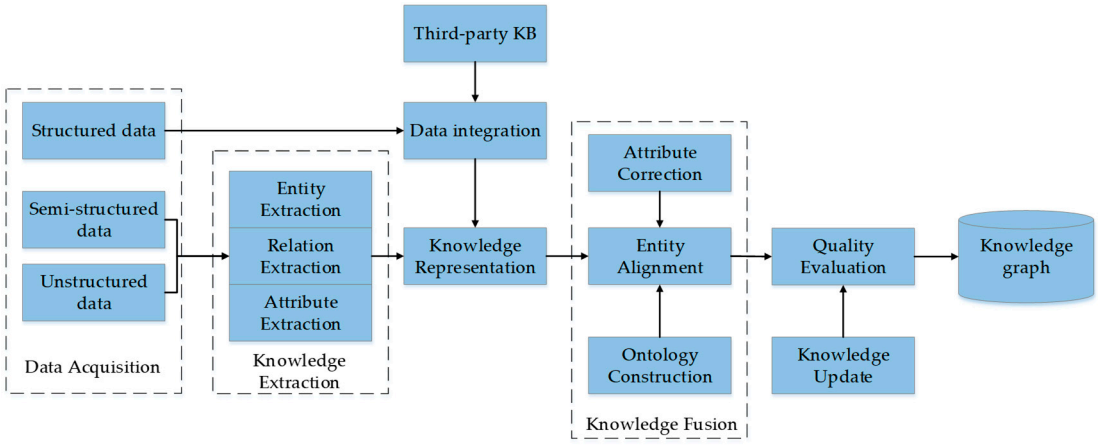


Figure 1. Construction framework of Chinese ancient historical and cultural knowledge graph.

3.1. Data Acquisition

According to the form of data storage, data sources can be divided into three categories, including structured data, semi-structured data, and unstructured data. In the process of constructing the knowledge graph of Chinese ancient history and culture, all data types in this paper include structured data (such as linked data, database), semi-structured data (such as tables and lists in web pages), and unstructured data (that is, plain text data), in which structured data comes from the data in the general domain encyclopedia knowledge graph on the Internet. The general domain knowledge graph that is used in this paper is Ownthink and CN-DBpedia. The semi-structured data mainly comes from Baidu Encyclopedia [40], HDWiki [41], taking zhishi.me as an example, Baidu Encyclopedia and Interactive Encyclopedia Chinese Network Encyclopedia are used as data sources for extracting a large amount of knowledge from them to build a knowledge graph [42], for which the semi-structured data are obtained through the use of wrappers, among which the generation methods of wrappers are divided into three categories: manual method, wrapper induction method, and automatic extraction method. In this paper, we mainly used the manual method to analyze and construct the rules of information extraction of the wrapper, so as to obtain the semi-structured data in the web page. The knowledge graph of ancient Chinese history and culture is to extract entities and related knowledge from semi-structured data (such as InfoBox) in Baidu Encyclopedia, as shown in Figure 2. For unstructured data, it comes from the data of ancient Chinese history related web pages on the Internet and the introduction part of Baidu Encyclopedia, we used the web crawler technology to obtain a large number of historical web page text data, and then used the obtained web page text data for processing [43].

The semi-structured data and unstructured data that were used in the construction of the ancient Chinese historical and cultural knowledge graph mainly come from the Internet, and the acquisition

method mainly uses the web crawler technology [44]. This paper implements a crawler that is based on the Python Scrapy framework [45] to obtain network data.

Figure 2. Baidu Encyclopedia data acquisition example (taking Nan Liang as an example), It is mainly divided into three parts. The first part is the Introduction part. This part is unstructured text data. After the extraction completed, the user named entity recognition and relationship extraction are performed. The second part is the Image part. This part mainly obtains the picture information. The obtained picture will be applied in the knowledge attribute query module of the knowledge graph system. The third part is the InfoBox part. This part is semi-structured data, which is mainly used to construct triples.

The process for obtaining semi-structured data is as follows: using web crawler technology according to the given initial web page (such as "中国历史朝代 (means: Chinese historical dynasty) link: https://baike.baidu.com/item/中国历史朝代/4056123"). A method that is similar to breadth-first traversal is used to crawl clickable page information in a webpage, and the obtained page information is saved as a html file in a webpage format. Subsequently, use the Xpath selector to extract the contents of the InfoBox in the saved web page. The main content is to save the contents of "basicInfo-item name" and "basicInfo-item value" in the InfoBox to a txt file. The same method saves the unstructured text content of the Introduction section to a txt file.

For unstructured text data, we view the web page format, write the corresponding crawling rules, scrape the desired data directly, and then save it to txt text for subsequent entity recognition and relationship extraction.

3.2. Knowledge Extraction

Knowledge extraction is the first step in the construction of knowledge graphs, which extracts structured information, such as entities, relationships, and entity attributes from semi-structured and unstructured data automatically [46]. The core technology involves entity extraction, relationship extraction, and attribute extraction.

1. **Named entity recognition:** Also known as entity extraction, refers to the automatic identification of named entities from text content, and it is the most critical part of information extraction. The quality of entity extraction will directly affect the work of relationship extraction and knowledge fusion in subsequent work. A deep learning algorithm is used for named entity recognition in this paper. A variety of methods are used for comparison experiments, and the

conditional random field is used as a benchmark. The BiLSTM, BiLSTM-CRF, and BiLSTM-CNN-CRF methods are used for comparison experiments on custom data sets. Finally, the BiLSTM-CNN-CRF model is used to extract the entities in the text. Specific experiments on named entity recognition will be explained in detail in the fourth part of the paper.

2. **Relationship extraction:** After obtaining the entities, the relationships between the entities need to be extracted from the relevant corpus, and the unrelated entities that are initially connected are connected through the relationship to form a knowledge network structure. This article will use the Chinese relation extraction tool DeepKE that was developed by Zhejiang University to extract the relationships between entities.
3. **Attribute extraction:** Attribute extraction is to aggregate the information of the same entity from various data to achieve the complete outline of the entity attributes. The attribute extraction in this article mainly comes from the semi-structured data in the information box in Baidu Encyclopedia. The knowledge extraction in this part mainly uses Python programming, using the Xpath selector to extract the contents of the InfoBox in the saved web page. Mainly save the contents of "basicInfo-item name" and "basicInfo-item value" in InfoBox to a txt file. The $ symbol is used to divide between "basicInfo-item name" and "basicInfo-item value". The information that is extracted from each webpage is saved to a txt file separately. Subsequently, the txt files that are extracted from multiple web pages are combined. It is finally stored as a triple. Figure 3 shows the results after processing.

Figure 3. Data extraction results from InfoBox in Baidu Encyclopedia. The left side of the figure is to extract the information in Infobox, which mainly includes the contents of "basicinfo item name" and "basicinfo item value" in the information box, and the $ symbol is used for segmentation in the middle. On the right side of the figure is the triple data result of information processing in the extracted Infobox, in the form of (entity1-relationship-entity2), in which the $ symbol is used to divide the entity and relationship.

3.3. Knowledge Fusion

After knowledge extraction is completed, the entities, relationships, and entity attributes are extracted from the obtained unstructured and semi-structured data. However, these results may contain other error information. Erroneous data must be eliminated in order to ensure the quality of knowledge [47]. Knowledge fusion includes two parts: entity alignment and attribute value filling. Among them, entity alignment is to determine whether there are multiple entities in the obtained entity

pointing to the same entity in the objective world [48]. The solution is to obtain knowledge from the third-party knowledge base Ownthink as input and supplement the entity attribute values.

3.4. Graph Construction

In this section, the processed structured data are mainly stored in the graph database, the database holding the triples data is the Neo4j graph database (https://neo4j.com/). The stored data include entities and entity relationships, entity attributes, and entity attribute values [47]. Cypher statements are used to store triples in the Neo4j graph database. By "MATCH (e: History), (cc: Attribute) WHERE e.Name = '% s' AND cc.Name = '% s' CREATE (e)-[r:% s {relation: '% s'}]-> (cc) RETURN r"statement to create a knowledge graph.

The following Figure 4 shows the storage transformation process.

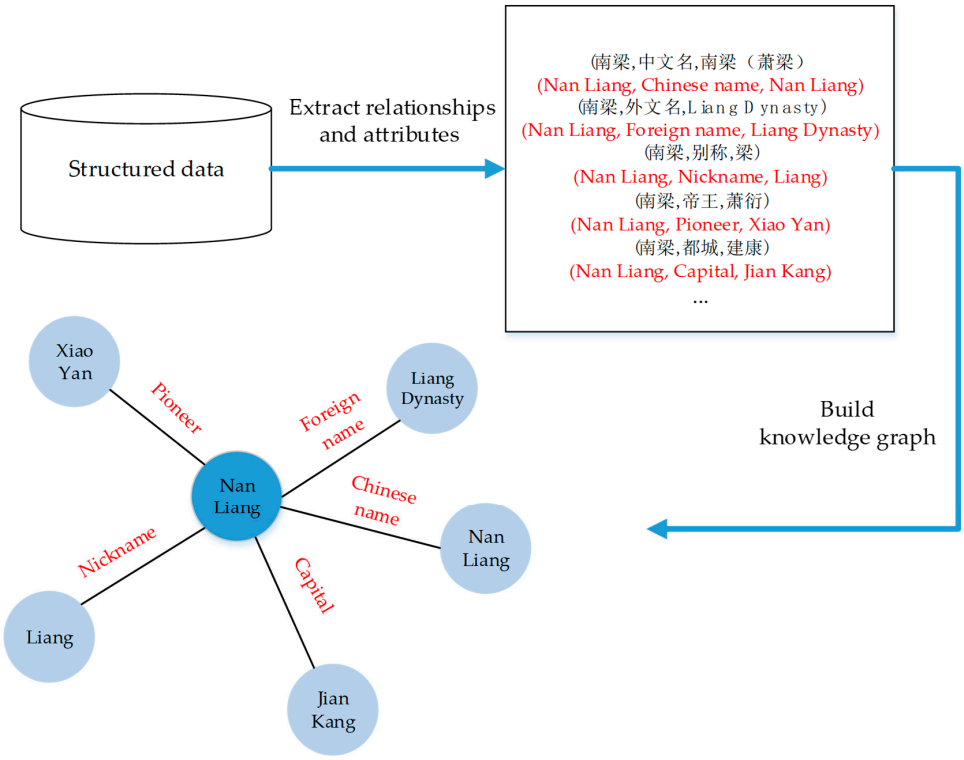


Figure 4. Structured data to graph database stored procedures.

Two principles in the construction of the knowledge graph:

1. The type of each node is represented by its identifier. If the identifier is "dynasty", the type of node created is "dynasty".
2. In the entity-relation table, the storage format of the triples data is "Entity1-Relation-Entity2" or "Entity1-Attribute-Attribute Value".

4. Acquisition of Chinese Ancient Historical and Cultural Knowledge

4.1. Overall Framework of the Model

Figure 5 shows the BiLSTM-CNN-CRF model framework. The BiLSTM-CNN-CRF model consists of three parts. The first part is CNN module, the second part is BiLSTM module, and the third part is CRF module. The CNN model process consists of three steps, training the word vector of the data set, extracting the character vector of the sentence by CNN network, and convoluting and maximizing the pool operation to obtain the character level characteristics of each word.

For the BiLSTM model, the character vector and the word vector are spliced, and the spliced word vector is used as input into the BiLSTM neural network model for entity recognition.

For the CRF model, the output of the BiLSTM model is decoded in order to get an optimal tag sequence.

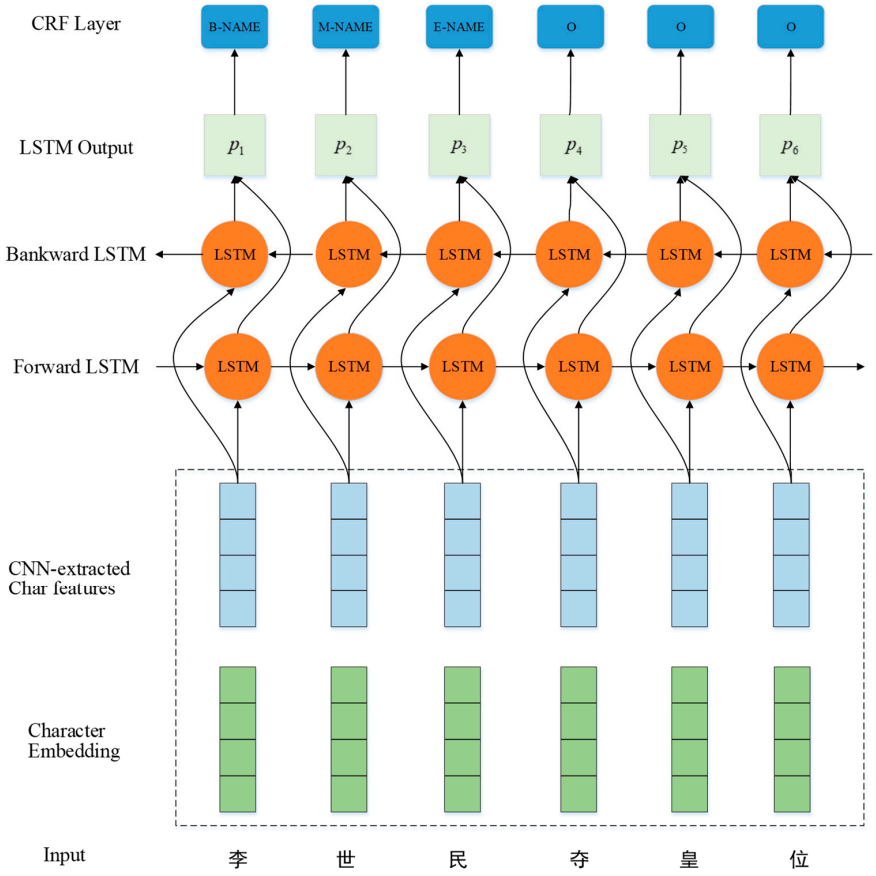


Figure 5. Bidirectional short-term memory-Convolutional Neural Networks-Conditional Random Field (BiLSTM-CNN-CRF) model of named entity recognition of Chinese ancient historical and cultural knowledge graph. The overall structure of the model is divided into three levels: (1) comprehensive embedding, including character embedding and word embedding. (2) Bidirectional short-term memory (BiLSTM) layer. This layer is used to capture historical and future information in the sentence of this article. (3) Conditional Random Field (CRF) layer. This layer is labeled with CRF. B in B-NAME indicates the beginning, NAME indicates the name of the person, M in M-NAME indicates the middle, NAME indicates the name of the person, E in E-NAME indicates the end, and NAME indicates the name of the person. The Chinese input is "李世民夺皇位" (means: Li Shimin seizes the throne), where Li Shimin is a person's name.

4.2. Feature Representation

The popular text representation usually uses the word bag model, because the model is simple to construct and it can reduce the complexity of vector calculation [49]. However, there are many shortcomings in this model. For example, the feature dimension of text will be very high when the sample data is large, which might lead to dimension explosion. The word vector matrix is very sparse, so it is easy to over fit. In order to solve this problem, Mikolov et al. [50] introduced the word embedding model, which is an effective method for learning high-quality word vector representation from a large number of unstructured text data, capturing very important syntactic and semantic information.

Currently, Word2vec [51] is the most widely used word vector training tool in the field of natural language processing, which mainly includes two training methods: Continuous Bag-of-Words Model

(CBOW) and Skip-gram. CBOW uses the context statement of a word as input to predict the current word, while Skip-gram uses the current word to predict the context statement around it. In this paper, the CBOW method is used to train word vectors for the data corpus acquired on the network, as shown in Figure 6.

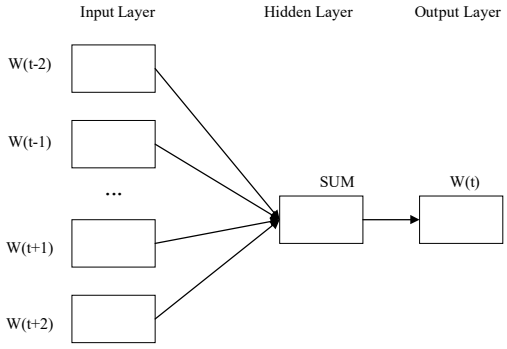


Figure 6. Training the word vectors using the Continuous Bag-of-Words Model (CBOW) mode, predict the current word from the input of the context sentence.

The optimization objective function of CBOW model in the training process is as follows:

$$L = \sum_{WeC} log\, p(W|Content(W)) \tag{1}$$

4.3. CNN Model

The convolutional layer in the convolutional neural network can extract the local feature information of the text data, and the most representative part of the local features can be further extracted as the feature vector through maximum pooling. Chiu, Nicholas et al. [52] used CNN in order to extract character-level features to achieve good results in the general field. Therefore, this paper uses CNN to extract the character-level features of words in Chinese ancient historical texts, and improves the model's performance by combining word vectors and character-level features. Figure 7 shows the CNN structure. It mainly includes a convolution layer, a pooling layer, and a character vector.

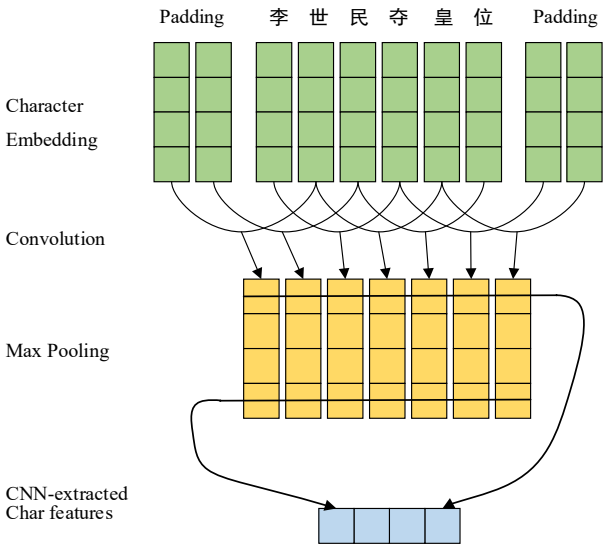


Figure 7. The convolution neural network for extracting character-level representations of words. Filled arrows indicate a dropout layer applied before character embeddings are input to CNN. The Chinese input is "李世民夺皇位" (means: Li Shimin seizes the throne), where Li Shimin is a person's name.

4.4. BiLSTM Model

Recurrent Neural Network (RNN) is an important kind of structure in neural networks. In theory, RNN can dynamically capture sequence data information. However, in actual use, it is prone to cause problems, such as gradient disappearance and gradient explosion. Long Short-Term Memory [53] can solve this kind of problem well. For named entity recognition, the positions of the entity sentences that need to be identified are different, and the importance of context information in the text is also different. This paper uses Bidirectional Long Short-Term Memory structure for model training in order to make better use of the information in the context.

4.4.1. LSTM Unit

LSTMs are variants of RNNs that are designed to cope with these gradient vanishing problems. Basically, a LSTM unit is composed of three multiplicative gates that control the proportions of information to forget and pass on to the next time step. These gating units are Input Gate, Forget Gate, and Output Gate. Figure 8 gives the basic structure of an LSTM unit.

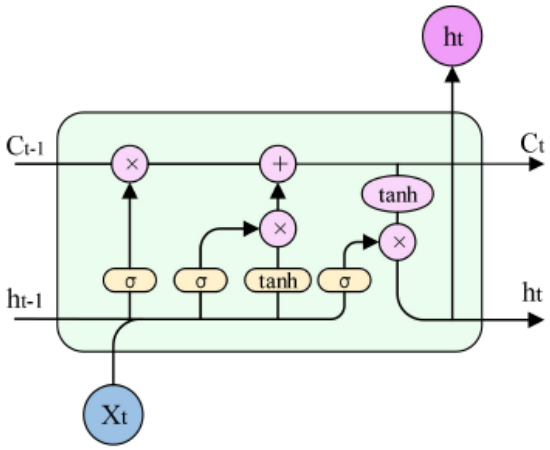


Figure 8. Schematic of LSTM unit.

Formally, the formulas to update an LSTM unit at time t are:

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \tag{2}$$

$$f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right) \tag{3}$$

$$\widetilde{c_t} = \tanh(W_c h_{t-1} + U_c x_t + b_c) \tag{4}$$

$$c_t = f_t \odot c_{t-1} + i \odot \widetilde{c_t} \tag{5}$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \tag{6}$$

$$h_t = o_t \odot \tanh(c_t) \tag{7}$$

where σ is the element-wise sigmoid function and $\odot$ is the element-wise product. x_t is the input vector (e.g., word embedding) at time t, and h_t is the hidden state (also called output) vector storing all of the useful information at (and before) time t. U_i, U_f, U_c, U_o denote the weight matrices of different gates for input x_t, and W_i, W_f, W_c, W_o are the weight matrices for hidden state h_t. b_i, b_f, b_c, b_o denote the bias vectors.

4.4.2. BiLSTM Unit

After using LSTM, the hidden layer state sequence with the same length as the sentence $[h_1, h_2 \ldots\ldots h_n]$ will be obtained. The BiLSTM network used in this paper can get the hidden layer

state in the forward and backward directions at time t. Sequences $[\overrightarrow{h_1}, \overrightarrow{h_2} \ldots\ldots \overrightarrow{h_n}]$ and $[\overleftarrow{h_1}, \overleftarrow{h_2} \ldots\ldots \overleftarrow{h_n}]$, the resulting hidden layer state sequences are forward and backward. Hidden layer state sequence stitching is generated, that is $h_t = [\overrightarrow{h_t} : \overleftarrow{h_t}]$. Figure 9 illustrates the basic structure of an BiLSTM unit.

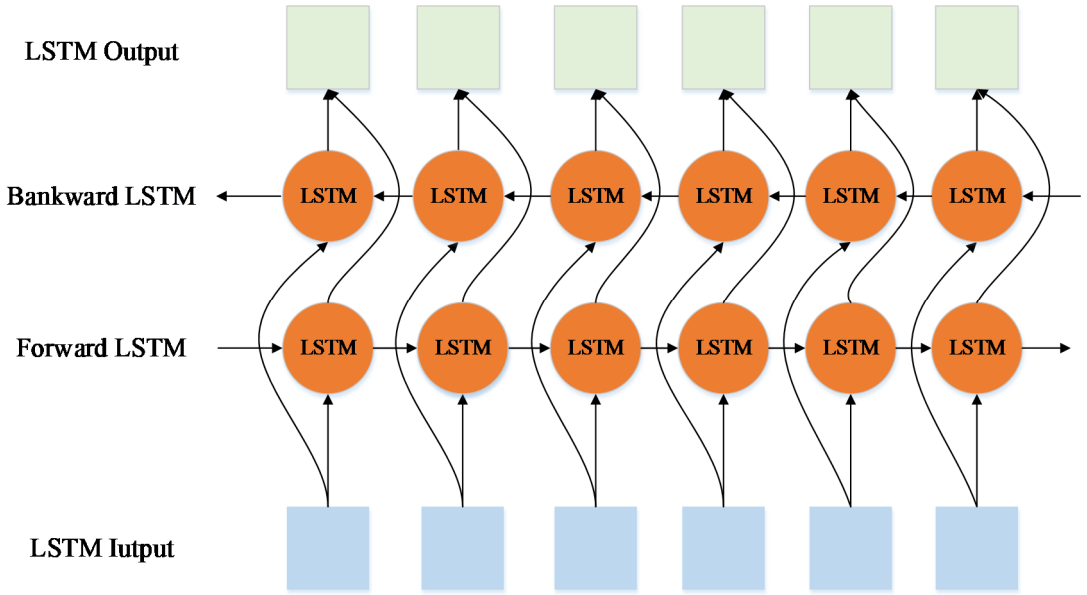


Figure 9. Bidirectional Long Short-Term Memory model.

4.5. CRF Model

Conditional random field is a probabilistic undirected graph model [54], and it is also a common algorithm in a sequence tagging task, which can be used for entity class tagging. In this paper, CRF layer is regarded as the last layer of neural network structure, and the output of BiLSTM module is processed in order to obtain the optimal global label sequence.

For a given text, $X = (x_1, x_2, x_3 \ldots x_n)$ is the input sentence, $y = (y_1, y_2, y_3 \ldots y_n)$ represents the output tag sequence, then the tag sequence score is as:

$$S(X, y) = \sum_{i=0}^{n} A_{y_i, y_{i+1}} + \sum_{i=1}^{n} P_{i, y_i} \tag{8}$$

Here, A is the transfer fraction matrix and $A_{i,j}$ is the fraction that is transferred from label i to label j.

All of the possible sequence paths to generate the probability distribution of output sequence y need to be normalized, as shown in formula (9):

$$P(y|X) = \frac{e^{S(X,y)}}{\sum\limits_{\widetilde{y} \in Y_X} e^{S(X,\widetilde{y})}} \tag{9}$$

In the training process, the logarithmic probability of the correct tag sequence is maximized, as shown in formula (10):

$$\log\left(P(y^*|X)\right) = S(X, y^*) - \log\left(\sum_{\widetilde{y} \in Y_X} e^{S(X,\widetilde{y})}\right) \tag{10}$$

where Y_X is the sequence of all possible tags for the input sentence X.

In the final decoding, the sequence with the highest predicted total score is selected as the optimal sequence, as shown in formula (11):

$$y^* = arg \max_{\widetilde{y} \epsilon Y_X} S(X, \widetilde{y}) \tag{11}$$

5. Experimental Results and Analysis

5.1. Data Preparation

The data are used in this paper to obtain the relevant text data on the Internet through the crawler due to the lack of relevant data about Chinese ancient history and culture in the data set published on the Internet, and then the acquired corpus has been segmented, to stop using words and other processing, and the corpus has been entity marked with the information of person name, place name, time, dynasty, war, system and so on.

Here, we used Baidu Encyclopedia to generate data sets. The Baidu Encyclopedia is more like Wikipedia, which contains text, tables, and pictures to describe, as an introduction to the entity, similar to the triplet provided by the public knowledge graph (such as CN-DBpedia, Ownthink) to represent entity relationship entity [55]. Therefore, we can obtain the triple information and unstructured text data of the entity at the same time by crawling Baidu Encyclopedia.

The specific process is as follows:

First of all, we grabbed hundreds of thousands of pages from Baidu Encyclopedia. These pages contain relevant data in Infobox and text data regarding the entity in Baidu Encyclopedia. After filtering, each page crawled can be regarded as an entity introduction. In this page, structured information about entities is provided in Infobox. For example, Chinese name, nickname, national leader, time, etc. After processing, we can obtain triple information about this entity, and finally to form triple. Subsequently, deep learning algorithm is used for unstructured text in the page to extract entities and relationships, and finally to constructs triples.

5.2. Data Annotations

Presently, the main annotation models of supervised learning include BIO, BIEO, BMESO, etc. The BMESO tagging method is used in the self-built dataset in this paper in order to be able to clearly represent the named entities to be recognized in the corpus. According to the research of Roth [56], Dai [57], Lample [33], BMESO is better than BIO, which can clearly divide the boundary of entities.

For each entity, the first word is marked as "B - (entity name)", the middle word as "M - (entity name)", the end word as "E - (entity name)", the single entity as "S- (entity name)", and the non-entity as "O". Table 1 shows the labeling strategy of BMESO

Table 1. Labeling strategy of BMESO.

Type	Start Tag	Middle Tag	End Tag
Time	B-TIME	M-TIME	E-TIME
Name	B-NAME	M-NAME	E-NAME
Location	B-LOC	M-LOC	E-LOC
Dynasty	B-DYN	M-DYN	E-DYN
Designation	B-CH	M-CH	E-CH
...	...	...	...
Non entity tag	O	O	O

Table 2 shows an example of entity annotation for a given Chinese ancient Chinese historical text while using the BMESO annotation strategy.

Table 2. An example of Chinese ancient historical text named entity annotation, Example "李渊于晋阳起兵 (means: Li Yuan starts his army in Jinyang)", where Li Yuan is a person's name and Jinyang is a place name.

English	**Li Yuan Starts His Army in Jinyang**						
Text	李	渊	于	晋	阳	起	兵
Tag	B-NAME	E-NAME	O	B-LOC	E-LOC	O	O

5.3. Evaluation Metrics

The standard evaluation measures like precision (P), recall (R), and F1-score ($F1$) are considered to evaluate our experiments.

$$P = \frac{TP}{TP + FP} \times 100\% \tag{12}$$

$$R = \frac{TP}{TP + FN} \times 100\% \tag{13}$$

$$F1\_score = \frac{2PR}{P + R} \times 100\% \tag{14}$$

5.4. Experimental Environment

The environment for all experiments is shown in Table 3:

Table 3. Experimental environment.

Project	**Environment**
Operating system	Ubuntu 16.04
GPU	NAVIDIA Quadro K1200
Hard disk	500G
Memory	8G
Python edition	3.6

5.5. Parameter Setting

The model that is proposed in this paper is built with Tensorflow framework, which is an in-depth learning framework that was developed by Google team and is the most widely used one among all frameworks at present. The experimental parameters are set as follows in Table 4:

Table 4. Experimental parameter setting.

Parameter Name	**Parameter Value**
Word vector dimension	300
Learning rate	0.001
Epoch	100
Dropout	0.5

5.6. Experimental Results and Analysis

CRF, BiLSTM, BiLSTM-CRF, and BiLSTM-CNN-CRF were used to test the custom data set to verify the effectiveness of the method used in this paper in order to verify the recognition effect of the method in this paper.

Experiment 1: Conditional random field (Baseline)

In this paper, the CRF model is used as the benchmark model to explore the performance of CRF model in self built history data set. During the experiment, CRF++ [58] toolkit, which is an open-source

tool of conditional random field, is used. The version of CRF + + 0.58 is used in the experiment. Entity recognition is realized by the building model.

During the experiment of CRF model, it is found that the model cannot correctly identify the names of people, places, and titles that never appear in the corpus. After analysis, the possible reasons are determined as follows: 1) the ancient surname is different from the modern surname, which might lead to a poor recognition effect; 2) most of the ancient place names have been changed, leading to mismatches; and, 3) at present, the data of people's names and place names marked in the language database is relatively small, leading to the model unable to correctly identify some entities.

The accuracy of CRF model is 75.45%, the recall rate is 72.38%, and the F1 value is 73.89%.

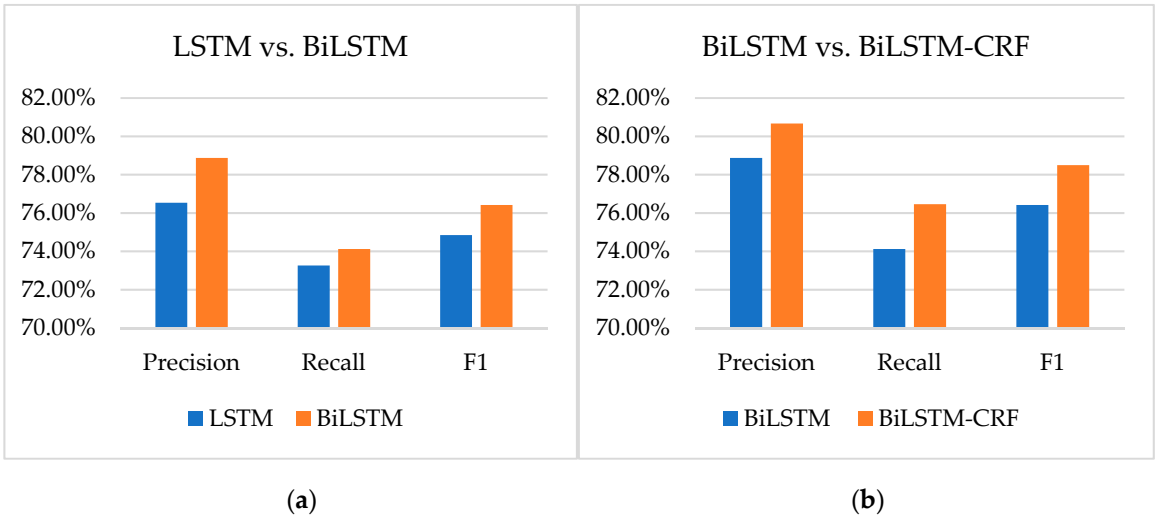
Experiment 2: LSTM vs. BiLSTM

Experiment 2 is mainly to verify the effectiveness of the LSTM bidirectional network structure. In this context, experiments on the LSTM model and the BiLSTM model were performed. The LSTM model can better capture the long-distance dependencies. However, there is still a problem in modeling sentences with LSTM. That is to say, it is impossible to encode information from back to front. However, the bidirectional LSTM combines the forward LSTM with the backward LSTM, so that both the front-to-back information and the back-to-front information can be encoded in the modeling process, thereby better modeling. According to the test results presented in Table 5, it can be seen that the entity recognition effect using the BiLSTM network model is better than the entity recognition while using the LSTM network structure. The main reason is that the BiLSTM network structure can make fuller use of context information than the LSTM network structure.

Table 5. Comparison of experimental results of LSTM model and BiLSTM model.

Method	Precision/%	Recall/%	F1_Score/%
LSTM	76.53	73.26	74.85
BiLSTM	78.87	74.12	76.42

The histogram of the experimental results of the LSTM model compared with the BiLSTM model is shown in (a) of Figure 10, below. It can be clearly seen that the BiLSTM is better than the LSTM in the results.



(**a**) (**b**)

Figure 10. Comparison of histograms of test results, (**a**) on the left, (**b**) on the right, (**a**) LSTM and BiLSTM test comparison histogram and (**b**) BILSTM and BiLSTM-CRF test result comparison.

Experiment 3: BiLSTM vs. BiLSTM-CRF

Experiment 3 is mainly to verify the performance of BiLSTM in the named entity recognition task on the custom dataset after adding CRF for decoding. The CRF layer can add some constraints

to the final constrained labels to ensure the validity of the predicted labels. These constraints are automatically learned by the CRF layer from the training data. The CRF layer decodes the output of the BiLSTM model to output the optimized marker sequence with the highest probability, so that the final prediction effect is better than that of the BiLSTM model without CRF. It can be seen from Table 6 that the effect of BiLSTM recognition after adding the CRF layer is better than that of decoding without the CRF layer, and the entity recognition performance is improved.

Table 6. Comparison of experimental results after adding CRF model.

Method	Precision/%	Recall/%	F1_Score/%
BiLSTM	78.87	74.12	76.42
BiLSTM-CRF	80.67	76.46	78.5

Figure 10b shows the histogram of the comparison results of BiLSTM and BiLSTM-CRF. In the results, BiLSTM-CRF is superior to BiLSTM in terms of accuracy, recall and F1 worth performance.

Experiment 4: BiLSTM-CRF vs. BiLSTM-CNN-CRF

Experiment 4 is mainly to verify the effectiveness of character level features extracted by CNN module in entity recognition. On the basis of Experiment 2, the experiments of BiLSTM-CRF and BiLSTM-CNN-CRF on self-built dataset are carried out. In the BiLSTM-CNN-CRF experiment, based on the BiLSTM-CRF experiment, the character level features that are extracted by CNN module are used, and then the extracted features and the trained word vectors are spliced, which are used as the input of BiLSTM model for training, and then the results of training are input into CRF model to select the optimal marking sequence through CRF, so as to get entity marking information. From the test results in Table 7, it can be seen that the entity recognition effect with CNN model added is due to that without CNN model added. From the experimental results, it can be concluded that CNN is helpful in entity recognition. Figure 11 shows the comparison histogram of two methods.

Table 7. Comparison of experimental results after adding CNN model.

Method	Precision/%	Recall/%	F1_Score/%
BiLSTM-CRF	80.67	76.46	78.5
BiLSTM-CNN-CRF	84.73	82.26	83.47

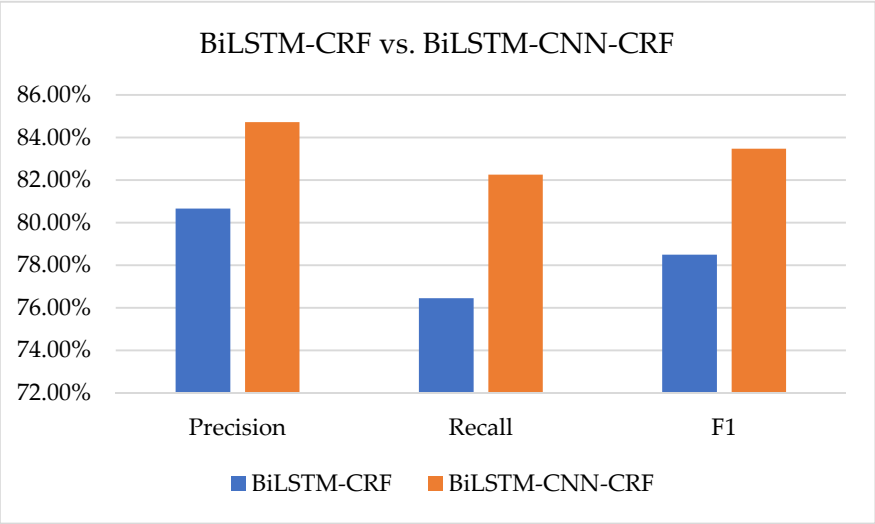


Figure 11. Histogram of BiLSTM-CRF and BiLSTM-CNN-CRF.

The above four experiments verified the validity of the CRF module, the CNN module, and the effectiveness of the BiLSTM network model.

BiLSTM-CRF is the leader in the traditional named entity recognition framework. In this model, the bidirectional long-term and short-term storage network solves the problem of dependencies between long-distance named entities in text. At the same time, the generated sequence is decoded due to the existence of the conditional random field, which further improves the recognition ability of the frame. Its overall effect is better than LSTM, BiLSTM.

BiLSTM-CNN-CRF is based on the BiLSTM-CRF and it adds a CNN layer. By adding the CNN layer to extract character-level features from the text data, the character-level features that are extracted through the CNN layer and the trained Word2vec are stitched together. The stitched vector is input to a Bidirectional long short-term memory network. After training, the output result is passed to the CRF, and the optimal sequence label is selected in the CRF layer.

Figure 12 shows the comparison histogram of the test results of different methods.

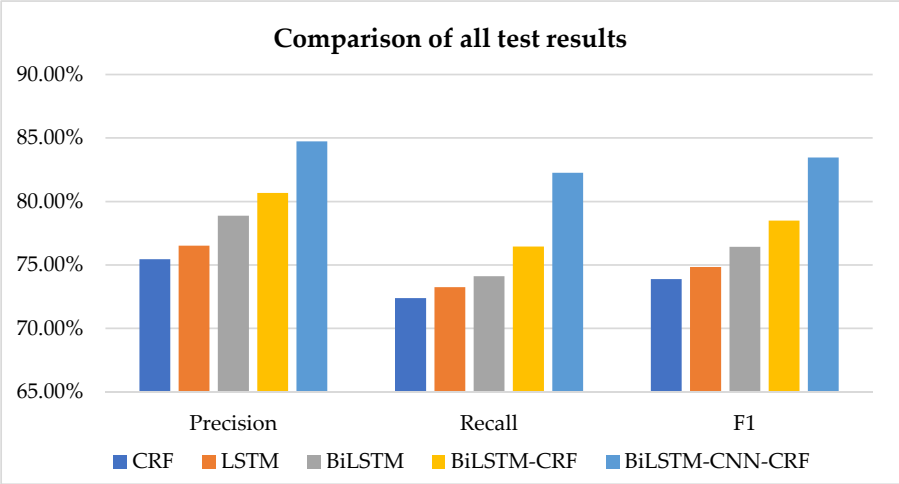


Figure 12. Histogram of comparative analysis of all experience results in this paper.

Table 8 shows the comparison of different test results.

Table 8. Comparison of experimental results of different methods.

Method	Precision/%	Recall/%	F1_Score/%
CRF	75.45	72.38	73.89
LSTM	76.53	73.26	74.85
BiLSTM	78.87	74.12	76.42
BiLSTM-CRF	80.67	76.46	78.5
BiLSTM-CNN-CRF	84.73	82.26	83.47

Based on the analysis of the above experimental results, the BiLSTM-CNN-CRF model that is used in this paper has achieved good results in the task of named entity recognition in the field of Chinese ancient history and culture.

6. Visual Display of Knowledge Graph

The above experiments can be utilized to obtain the triple data of Chinese ancient historical and cultural knowledge graph. In this paper, the entities and relationships of the acquired Chinese ancient historical and cultural knowledge graph are stored in neo4j graph database, with a total number of entities of about 15,000 and a total of 30 types of relationships. Using Echarts (similar to D3.js) in combination with the Flask framework, a knowledge graph system that was based on ancient

Chinese history and culture was developed, and the acquired ancient Chinese history and culture knowledge was visually displayed in various forms, such as text, pictures, and force-oriented diagrams. The system functions mainly include the following two parts:

1. Inquiry module of knowledge graph of ancient Chinese history and culture. The dynasty is used as an entity to expand and display historical figures of the same dynasty, the start and end time of the dynasty, and the name of the emperor. The knowledge graph supports zooming in, zooming out, and moving. The entities that have a relationship with the clicked entity can be highlighted when you click an entity with the mouse, while other unrelated entities are displayed in grayscale, as shown in Figure 13.
2. Entity attribute knowledge query module. This module is mainly divided into two parts for display: The first part is to display the relevant attribute knowledge of search entities while using force-oriented diagrams. The second part shows the picture information of this entity, and some information introduced by encyclopedia, as shown in Figure 14.

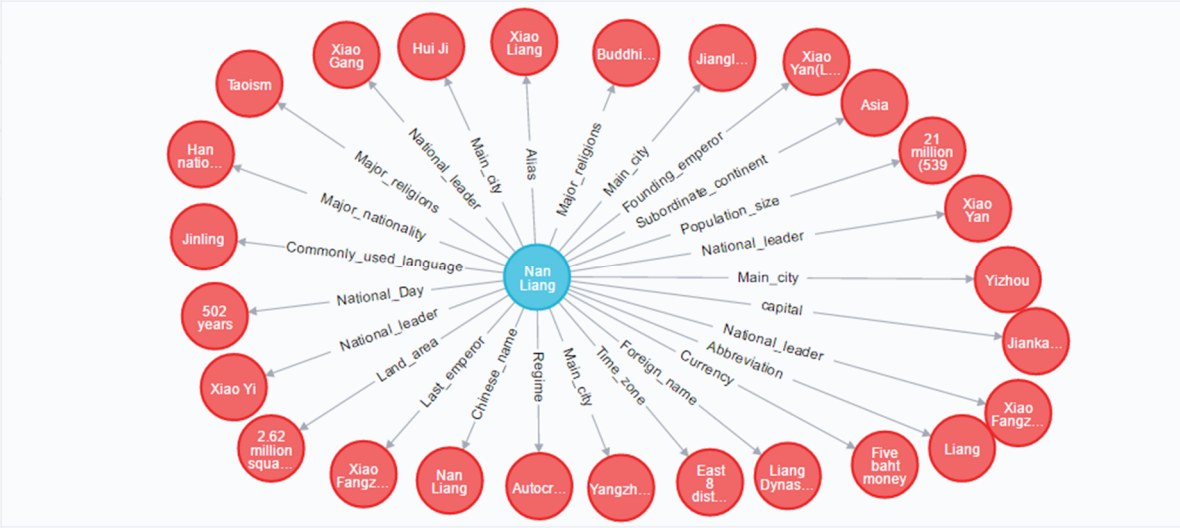


Figure 13. Chinese ancient history and culture knowledge graph query module. The blue circle is entity 1 and the red circle is entity 2. The text on the arrow indicates the relationship between the two.

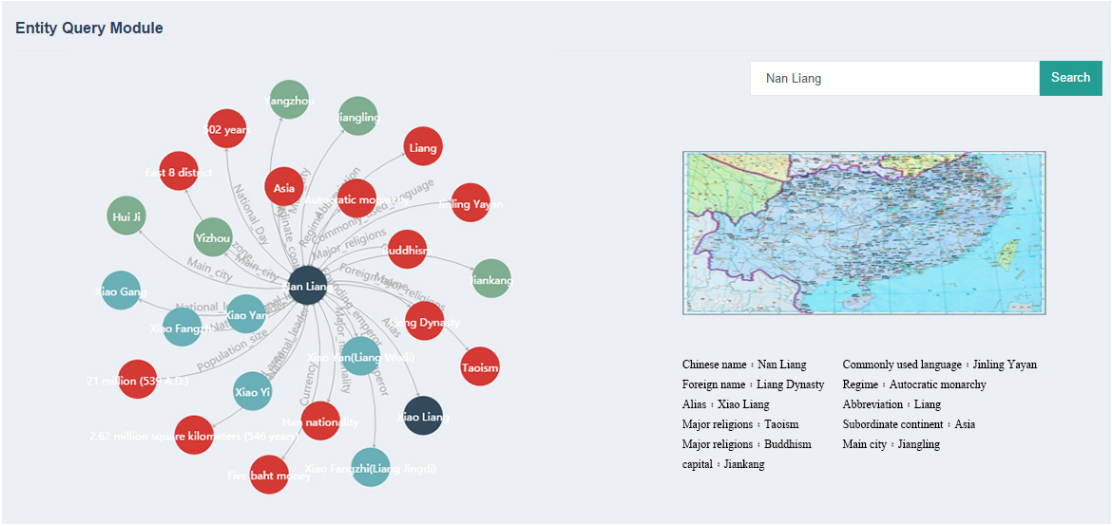


Figure 14. Entity attribute knowledge query module, the left is the force-oriented diagram of the query entity, and the right is the encyclopedia introduction of the query entity.

7. Conclusions

This paper put forward a construction process of Chinese ancient historical and cultural knowledge graph. Firstly, the data acquisition is introduced, which includes structured data, semi-structured data, and unstructured data. Then how to extract knowledge from unstructured text data is given and a BiLSTM-CNN-CRF neural network model for entity extraction is proposed. According to the experimental results, BiLSTM-CNN-CRF can extract entities from unstructured text better. Finally, the knowledge graph triplet is constructed and stored in the neo4j database that is based on the entity relationship, which is visualized by using Echarts and web programming.

In the future work, we will try to use Bert model to extract entity relationship in unstructured text, and further improving the effect of named entity recognition. At the same time, because the current relationship extraction part uses the open source tool DeepKE to extract relationships, the focus will be on the relationship extraction between entities in the future in order to improve the accuracy of relationship extraction.

The construction of the ancient Chinese historical knowledge graph has only just begun. In the future, we will work hard to build a large-scale and high-quality ancient Chinese historical knowledge graph.

Author Contributions: Writing—original draft, S.L.; Software, H.Y.; Visualization, J.L.; Validation, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant no. 61876031), Natural Science Foundation of Liaoning Province, China (grant no. 20180550921 and 2019-ZD-0175) and Scientific Research Fund Project of the Education Department of Liaoning Province (LJYT201906).

Acknowledgments: The authors would like to thank all anonymous reviewers and editors for their helpful suggestions for the improvement of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. China Internet Information Center. CNNIC: Statistical Report on Internet Development in China in 2019. Available online: http://www.cnnic.net.cn/ (accessed on 30 March 2020).
2. Singhal, A. Introducing the Knowledge Graph: Things, Not Strings. Available online: https://googleblog.blogspot.com/2012/05/introducing-knowledge-graphthings-not.html (accessed on 15 October 2019).
3. Biega, J.; Kuzey, E.; Suchanek, F.M. Inside YAGO2s: A transparent information extraction architecture. In Proceedings of the 22nd International Conference on World Wide Web Companion, Rio de Janeiro, Brazil, 13–17 May 2013; International World Wide Web Conferences Steering Committee: Rio de Janeiro, Brazil, 2013; pp. 325–328.
4. Bizer, C.; Lehmann, J.; Kobilarov, G.; Auer, S.; Becker, C.; Cyganiak, R.; Hellmann, S. DBpedia-A crystallization point for the Web of data. *J. Web Semant.* **2009**, *7*, 154–165. [CrossRef]
5. Erxleben, F.; Günther, M.; Krötzsch, M.; Mendez, J.; Vrandečić, D. Introducing wikidata to the linked data web. In Proceedings of the 13th International Semantic Web Conference, Riva del Garda, Italy, 19–23 October 2014.
6. Xu, B.; Xu, Y.; Liang, J.; Xie, C.; Liang, B.; Cui, W.; Xiao, Y. CN-DBpedia: A Never-Ending Chinese Knowledge Extraction System. In Proceedings of the International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Arras, France, 27–30 June 2017; Springer: Cham, Switzerland, 2017; pp. 428–438.
7. Niu, X.; Sun, X.; Wang, H.; Rong, S.; Qi, G.; Yu, Y. Zhishi. me-weaving chinese linking open data. In Proceedings of the Semantic Web–ISWC 2011, Bonn, Germany, 23–27 October 2011; Springer: Berlin, Germany, 2011; pp. 205–220.
8. MrYener. OwnThink Knowledge Graph. Available online: https://www.ownthink.com/ (accessed on 30 March 2020).
9. Wang, Z.; Li, J.; Wang, Z.; Li, S.; Li, M.; Zhang, D.; Shi, Y.; Liu, Y.; Zhang, P.; Tang, J. XLore: A Large-scale English-Chinese Bilingual Knowledge Graph. Presented at the Meeting of the International Semantic Web Conference (Posters & Demos), Sydney, Australia, 21–25 October 2013.

10. IMDB Official. IMDB. Available online: http://www.imdb.com (accessed on 30 March 2020).
11. MetaBrainz Foundation. Musicbrainz. Available online: http://musicbrainz.org/ (accessed on 30 March 2020).
12. Knowledge Map of Traditional Chinese Medicine. Available online: http://www.tcmkb.cn/kg/index.php (accessed on 30 March 2020).
13. Audema; Yang, Y.; Sui, Z.; Dai, D.; Chang, B.; Li, S.; Xi, H. Preliminary Study on Construction of Chinese Medical Knowledge Atlas CMeKG. *J. Chin. Inf. Process.* **2019**, *33*, 1–9.
14. Pandolfo, L. "STOLE: A Reference Ontology for Historical Research Documents." DC@ AI* IA. 2015. Available online: https://www.semanticscholar.org/paper/STOLE%3A-A-Reference-Ontology-for-Historical-Research-Pandolfo/90441c6089e278045980777a2fefb8fe5d41a41c (accessed on 30 March 2020).
15. Doerr, M. Ontologies for Cultural Heritage. In *Handbook on Ontologies*; Springer: Berlin/Heidelberg, Germany, 2009.
16. Wang, W.W.; Wang, Z.G.; Pan, L.M.; Liu, Y.; Zhang, J.T. Construction and Implementation of Historical Graph Knowledge Graph in Big Data Environment. *J. Syst. Simul.* **2016**, *28*, 2560–2566.
17. Gene Ontology Consortium. Available online: http://geneontology.org/ (accessed on 30 March 2020).
18. Hu, C.M.; Cai, W.C.; Huang, L.J.; Chao, C.J.; Hsu, C.Y. A nutrition analysis system based on recipe ontology. *Univ. Taipei Med.* **2006**, *15*, 57–71. [CrossRef]
19. Ruan, T.; Wang, M.; Sun, J.; Wang, T.; Zeng, L.; Yin, Y.; Gao, J. An automatic approach for constructing a knowledge base of symptoms in Chinese. In Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China, 15–18 December 2016; Volume 8, p. 33.
20. Wang, W.; Wang, Z.; Pan, L.; Liu, Y.; Zhang, J. Research on the Construction of Bilingual Movie Knowledge Map. *J. Peking Univ. (Nat. Sci. Ed.)* **2016**, *52*, 25–34.
21. Breast Cancer Knowledge Atlas. Available online: http://wasp.cs.vu.nl/BreastCancerKG/ (accessed on 30 March 2020).
22. Chi, Y.; Yu, C.; Qi, X.; Xu, H. Knowledge Management in Healthcare Sustainability: A Smart Healthy Diet Assistant in Traditional Chinese Medicine Culture. *Sustainability* **2018**, *10*, 4197. [CrossRef]
23. Huang, L.; Yu, C.; Chi, Y.; Qi, X.; Xu, H. Towards Smart Healthcare Management Based on Knowledge Graph Technology. In Proceedings of the 2019 8th International Conference on Software and Computer Applications, Penang, Malaysia, 19–21 February 2019; pp. 330–337. [CrossRef]
24. Haihong, E.; Zhang, W.J.; Xiao, S.Q.; Cheng, R.; Hu, Y.X.; Zhou, X.S.; Niu, P.Q. Survey of entity relationship extraction based on deep learning. *Ruan Jian Xue Bao/J. Softw.* **2019**, *30*, 1793–1818. (In Chinese). Available online: http://www.jos.org.cn/1000-9825/5817.htm (accessed on 30 March 2020).
25. Han, X.; Huang, D. Study of Chinese Part-of-Speech Tagging Based on Semi-Supervised Hidden Markov Model. *Small Microcomput. Syst.* **2015**, *36*, 2813–2816.
26. Borthwick, A.E. A Maximum Entropy Approach to Named Entity Recognition. Ph.D. Thesis, New York University, New York, NY, USA, 1999.
27. Wallach, H.M. Conditional Random Fields: An Introduction. *Tech. Rep.* **2004**, *53*, 267–272.
28. He, Y.; Luo, C.; Hu, B. A Geographic Named Entity Recognition Method Based on the Combination of CRF and Rules. *Comput. Appl. Softw.* **2015**, *32*, 179–185.
29. Wang, Z.; Jiang, M.; Gao, J.; Chen, Y. A Chinese Named Entity Recognition Method Based on BERT. *Comput. Sci.* **2019**, *46*, 138–142.
30. Collobert, R.; Weston, J.; Bottou, L.; Karlen, M.; Kavukcuoglu, K.; Kuksa, P. Natural Language Processing (Almost) from Scratch. *J. Mach. Learn. Res.* **2011**, *12*, 2493–2537.
31. Huang, Z.; Xu, W.; Yu, K. Bidirectional LSTM-CRF models for sequence tagging. *arXiv* **2015**, arXiv:1508.01991.
32. Santos, C.N.; Guimaraes, V. Boosting named entity recognition with neural character embeddings. *arXiv* **2015**, arXiv:1505.05008.
33. Lample, G.; Ballesteros, M.; Subramanian, S.; Kawakami, K.; Dyer, C. Neural architectures for named entity recognition. *arXiv* **2016**, arXiv:1603.01360.
34. Strubell, E.; Verga, P.; Belanger, D.; McCallum, A. Fast and accurate entity recognition with iterated dilated convolutions. *arXiv* **2017**, arXiv:1702.02098.
35. Feng, Y.H.; Yu, H.; Sun, G.; Sun, J.J. Named Entity Recognition Method Based on BLSTM. *Comput. Sci.* **2018**.
36. Maimai, A.; Wushou, S.; Palidan, M.; Yang, W. Uighur named entity recognition based on BILSTM-CNN-CRF model. *Comput. Eng.* **2018**, *44*, 230–236.

37. Li, L.S.; Guo, Y. Biomedical named entity recognition based on CNN-BILSTM-CRF model. *Chin. J. Inf.* **2018**, *32*, 116–122.
38. Yang, S.; Han, R. Method and Tool Analysis of KnowledgeMapping Abroad. *Libr. Inf. Knowl.* **2012**, *6*, 101–109.
39. Börner, K.; Chen, C.; Boyack, K.W. Visualizing knowledge domains. *Annu. Rev. Inf. Sci. Technol.* **2003**, *37*, 179–255. [CrossRef]
40. Baidu Encyclopedia the World's Largest Chinese Encyclopedia. Available online: https://baike.baidu.com/ (accessed on 30 March 2020).
41. HDWiki—More Authoritative Encyclopedia. Available online: http://www.baike.com/ (accessed on 30 March 2020).
42. Wu, T.; Qi, G.; Li, C.; Wang, M. A Survey of Techniques for Constructing Chinese Knowledge Graphs and Their Applications. *Sustainability* **2018**, *10*, 3245. [CrossRef]
43. Wang, H.; Qi, G.; Chen, H. *Knowledge Atlas: Method, Practice and Application*; Electronic Industry Press: Beijing, China, 2019; pp. 154–180.
44. Wang, H.; Fang, Z.; Zhang, L.; Pan, J.Z.; Ruan, T. Effective Online Knowledge Graph Fusion. In Proceedings of the Semantic Web-ISWC 2015, Bethlehem, PA, USA, 11–15 October 2015; Springer International Publishing: Bethlehem, PA, USA, 2015; pp. 286–302.
45. Tarjan, R.E. Finding optimum branchings. *Networks* **1977**, *7*, 25–35. [CrossRef]
46. Cowie, J.; Lehnert, W. Information extraction. *Commun. ACM* **1996**, *39*, 80–91. [CrossRef]
47. Wang, N.; Haihong, E.; Song, M.; Wang, Y. Construction Method of Domain Knowledge Graph Based on Big Data-Driven. In Proceedings of the 2019 5th International Conference on Information Management (ICIM), Cambridge, UK, 24–27 March 2019; pp. 165–172. [CrossRef]
48. Huang, H.; Yu, J.; Liao, X.; Xi, Y. Summary of Knowledge Graph Research. *Appl. Comput. Syst.* **2019**, *28*, 1–12. Available online: http://www.c-s-a.org.cn/1003-3254/6915.html (accessed on 30 March 2020).
49. Meng, J.; Long, Y.; Yu, Y.; Zhao, D.; Liu, S. Cross-Domain Text Sentiment Analysis Based on CNN_FT Method. *Information* **2019**, *10*, 162. [CrossRef]
50. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed representations of words and phrasesand their compositionality. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 3111–3119.
51. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
52. Chiu, J.P.C.; Nichols, E. Named entity recognition with Bidirectional LSTM-CNNs. *Trans. Assoc. Comput. Linguist.* **2016**, *4*, 357–370. [CrossRef]
53. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]
54. Wu, H.; Lu, L.; Yu, B. Chinese Named Entity Recognition Based on Transfer Learning and BiLSTM-CRF. *Small Micro Comput. Syst.* **2019**, *40*, 1142–1147.
55. Han, X.; Zhang, Y.; Zhang, W.; Huang, T. An Attention-Based Model Using Character Composition of Entities in Chinese Relation Extraction. *Information* **2020**, *11*, 79. [CrossRef]
56. Ratinov, L.; Roth, D. Design Challenges and Misconceptions in Named Entity Recognition. In Proceedings of the Thirteenth Conference on Computational Natural Language Learning, Boulder, CO, USA, 4–5 June 2009.
57. Dai, H.J.; Lai, P.T.; Chang, Y.C.; Tsai RT, H. Enhancing of chemical compound and drug name recognition using representative tag scheme and fine-grained tokenization. *J. Cheminform.* **2015**, *7* (Suppl. 1), S14. [CrossRef] [PubMed]
58. CRF++/Wiki/Home. Available online: https://sourceforge.net/p/crfpp/wiki/Home/ (accessed on 30 March 2020).