Journal of Network Intelligence ©2022 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **7**, Number **2**, May **2022**

Chinese Text Entailment Recognition Considering Characteristic Vocabulary and Negative Vocabulary

Jingdong Wang
School of computer science
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
School of information engineering
Guandong Atv Academy For Performing Arts
Hujing Road, Dongguan City,Guangdong Province, 523710, China
wangjingdong@neepu.edu.cn

Huimin Li
School of computer science
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
1363022737@qq.com

Fanqi Meng*
School of computer science
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
School of information engineering
Guandong Atv Academy For Performing Arts
Hujing Road, Dongguan City,Guangdong Province, 523710, China
*Corresponding Author: mfq81@163.com

Yujie Zheng
School of computer science
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
278764886@qq.com

Peifang Wang
School of computer science
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
w13578519072@163.com

Xiaolong Yang
School of Economics and Management
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
yangxiaolong@neepu.edu.cn

Jieping Han
School of Economics and Management
Northeast Electric Power University
Changchun Road, Jilin City, Jilin Province, 132012, China
hanjieping@126.com

Received November 2021; revised January 2022

Abstract. *Accurate recognition of text entailment can effectively support artificial intelligence tasks such as paraphrase, abstract, translation, and dialogue. However, the accuracy of Chinese text entailment recognition still needs to be improved, particularly the appearance of negative words, it will further reduce the accuracy of Chinese text entailment recognition. In order to improve the accuracy, a Chinese text entailment recognition method considering feature vocabulary and negative vocabulary is proposed. This method has 4 steps: the first step is to extract the characteristic vocabulary (noun, verb, adjective, adverb) and negative vocabulary of the premise sentence and hypothesis sentence respectively, enrich the text semantics, and then encode sentences and vocabulary; the second step is to use dual-tower Bi-LSTM and CNN to extract the semantic features of the text in the encoding; the third step is to input the semantic features of the text into the fully connected layer to realize the semantic matching of the premise sentence and the hypothesis sentence, then initially identify the entailment relationship; the fourth step is to use the correction layer to repeatedly correct the entailment relationship until the best effect is achieved. The experimental results show that the accuracy of this method is about 5% higher than that of the traditional method. It lays the foundation for further efficient recognition of text entailment with negative words.*
Keywords: text entailment, Bi-LSTM, semantic matching, CNN, negative vocabulary

1. **Introduction.** Text entailment is to determine the relationship between two sentences given two sentences, which is an important basis for tasks such as machine reading comprehension, dialogue systems, and information retrieval. At present, Chinese text entailment methods are mainly divided into two categories: traditional recognition methods and recognition methods based on machine learning. The traditional entailment recognition method is not suitable for text with complex structure; the method based on machine learning is difficult to completely extract the deep features of the sentence and the model has a high dimensionality. In recent years, with the development of text entailment tasks, there are more and more texts containing negative words. However, because the existing methods are difficult to understand the advanced grammar of texts well, the recognition effect of text entailment containing negative words is not good [1–3].

Based on the above analysis, in order to better recognize the entailment relationship of Chinese texts containing negative words, a Chinese text entailment recognition method that considers feature words and negative words is proposed. This method innovatively extracts feature vocabulary and negative vocabulary and trains it together with the text, and uses the dual-tower model to train the premise sentence and hypothesis sentence respectively, reducing the model dimensionality, and solving the problem of single text feature and long training time in traditional methods; on this basis, a correction layer is added to effectively solve the problem of low accuracy of traditional machine learning methods for text recognition with negative words.

The main contribution of this article:

(1) Using Bi-LSTM and CNN to extract the semantic information of the text twice, the contextual semantic features and local semantic features of the text can be extracted, which solves the problem of incomplete semantic extraction in traditional methods.

(2) It is proposed to extract feature vocabulary and negative vocabulary to enrich the semantics in the method input stage, which solves the problem of low accuracy of the traditional model for text recognition with negative vocabulary, and at the same time makes effective explorations for semantic enrichment.

(3) In the method modification stage, the negative difference set and negative vocabulary modification method parameters are added to provide a reference method for the text entailment recognition task with negative vocabulary.

(4) In the special Chinese text entailment recognition task with negative vocabulary, a large number of comparisons of the accuracy of the mainstream text entailment model in the special Chinese text entailment provide a reliable reference method for the research of special text matching, special text similarity and other tasks.

The remaining parts are organized as follows: Section 2 reviews related work; Section 3 describes Chinese text entailment recognition methods that consider characteristic vocabulary and negative vocabulary; Section 4 verifies the effectiveness and stability of the method through experiments; Section 5 summarizes the work done, and tells the shortcomings and future research directions.

2. **Related Work.** Recognizing text entailment is an important and complex task, involving techniques such as grammatical rules, logical deduction, and semantic reasoning. Currently, text entailment recognition mainly includes two methods: traditional recognition methods and recognition methods based on machine learning.

Traditional recognition methods are mainly divided into text entailment recognition methods based on similarity and text entailment recognition methods based on logical calculus. The similarity-based method is relatively simple and intuitive, and was once the most popular method for text entailment recognition. Jijkoun and Rijke [4] proposed the use of bag-of-words model for textual entailment recognition by first slicing the lexical language of the text, by lexical frequency based on different single-word weights, and then calculating Lin-similarity [5] and WordNet similarity [6], and finally using these two similarities for textual entailment recognition. Kouylekov and Negri [7] improved the distance algorithm by first proposing a similarity calculation rule based on the grammar tree, proposing that if the grammar tree of the predicate sentence can be walked to the grammar tree of the hypothesis sentence by branching of the tree then the predicate sentence and the hypothesis sentence are Entailment relations. Heilman and Smith [8] enriched the syntactic tree editing rules based on Mehdad with new definitions for the operations of adding, deleting and changing major nodes in the tree, which finally led to a substantial improvement in the text recognition accuracy. The identification method based on logical algorithms is closely related to the deductive reasoning of the mathematical community, and this method is more complex. Hobbs et al. [9] first applied the method of Abductive Reasoning for textual reasoning. Attempting to prove the relationship between two texts using rigorous reasoning in the mathematical community, they constructed a large collection of common knowledge as well as facts specific to certain domains, creating a large-scale database for posterity, but this method implies recognition accuracy is not very high. In fact, transforming natural language into mathematical language according to the transformation rules, there is a large amount of distortion in between. To reduce the degree of distortion Bos and Johan [10] introduced a relaxation strategy to the field of textual implication recognition. Similarly, Moldovan et al. [11] proposed to allow ignoring some sub-propositions of the text without affecting the confidence of the premise and hypothesis sentences, and also to introduce fuzzy inference rules to transform natural language into mathematical expressions.

With the rapid development of machine learning, Raina et al. [12] have proposed text entailment recognition methods based on machine learning. Among them, the entailment recognition method based on neural network and the recognition method based on pre-training are the most popular. The method based on neural network has effectively improved the accuracy of text entailment recognition in recent years. Kim [13] and Kalchbrenner et al. [14] and others proposed for the first time the use of convolutional neural networks in the field of natural language processing. MouL et al. [15] first proposed the use of convolutional neural networks in the field of text entailment recognition, using convolutional neural networks to encode text and extract semantic information, then use softmax to classify text entailment. This is a major breakthrough, which greatly improves the accuracy of text entailment recognition. However, because CNN is not sensitive to the order in which text appears, Bowman et al. [16] introduced Long Short-Term Memory Network in the study of text entailment recognition, so that the model can effectively learn the order of the text, the output state of the LSTM is the semantic feature of the text, and because the gate mechanism of LSTM can control the disappearance of the gradient, the effect of using LSTM for text entailment has always been better. On this basis, Bowman et al. proposed a text entailment recognition method based on twin LSTM [17], since the input text is in pairs, it fits well with the Siamese network, but this method can only extract semantic latent features and is not suitable for entailment recognition of complex long texts. With the explosion of pre-training models, the BERT model [18] is a classic pre-training model, which was used in the field of text entailment recognition when it was proposed. BERT does a lot of pre-training on the text inside, and only needs to fine-tune the parameters when doing fine-turn.

Traditional recognition methods are difficult to extract the deep-level semantic information of the text, and waste a lot of text semantic information, so the accuracy of text entailment recognition is limited. In the text entailment recognition method based on deep learning, usually due to problems such as large dimensions and many parameters, the model training speed is slow and requires good hardware support. Because the above two methods are difficult to better understand the advanced grammar of the text (for example: double negation means affirmation), the entailment recognition effect of special texts containing negative words is not very good. To this end, a Chinese text entailment recognition method considering feature vocabulary and negative vocabulary is proposed, extract the characteristic vocabulary and negative vocabulary of the text, and input them into the Bi-LSTM+CNN model together with the corresponding sentence. Since Bi-LSTM is good at learning the semantic information of the text context, and CNN is good at learning the local information of the text, Bi-LSTM and CNN are combined. Taking into account the problems of traditional machine learning-based methods such as large dimensions, many parameters, and slow training, the dual-tower structure is used to train the premise sentence and hypothesis sentence separately, the dimensionality of the method is greatly reduced, which greatly reduces the training time of the method.

3. **Chinese text entailment recognition method considering characteristic vocabulary and negative vocabulary.** The overall framework of the Chinese text entailment recognition method considering characteristic vocabulary and negative vocabulary is shown in Figure 1. The method generally includes four steps: text preprocessing, semantic feature extraction, preliminary recognition of text entailment, and result correction.

In the text preprocessing stage, clear deactivated words, digital normalization and extraction of feature words and negative words for the premise sentence and hypothesis sentence. Then Bi-LSTM and CNN are used to extract the semantic features of the text in two layers; then the text entailment is initially recognized through the fully connected layer. Finally, the entailment result of the preliminary recognition is corrected in the result correction stage until a better recognition result is obtained.

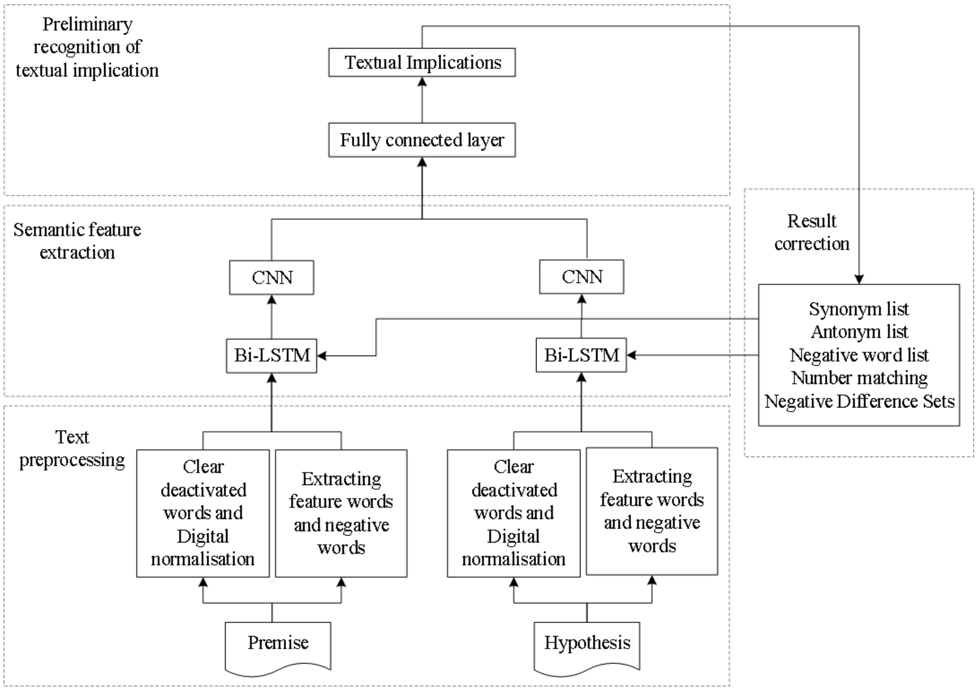

Figure 1. The overall framework of the Chinese text entailment recognition method considering feature vocabulary and negative vocabulary

3.1. **Text preprocessing.** Aiming at the problems of large dimensions and long training time of current neural network models, a dual-tower structure is proposed to train the premise sentence and hypothesis sentence separately. First, preprocess the premise sentence and hypothesis sentence and extract the characteristic vocabulary and negative words. Preprocessing mainly includes clear deactivated words and digital unification; research shows that the four types of words, nouns, verbs, adjectives, and adverbs, contain the most semantically rich words [19], so the characteristic vocabulary belongs to these four types of words.

(1)Clear deactivated words

Structural features usually cause excessive text noise. First, the text must be clear deactivated words. Deactivated words refer to words that have no actual semantics. The clear deactivated words collection can achieve the effect of data enhancement. This article compares the "Chinese deactivated vocabulary list", "Harbin Institute of Technology deactivated vocabulary list", "Baidu deactivated vocabulary list" and "Sichuan University Machine Intelligence Laboratory deactivated vocabulary" (*https://github.com/goto456/stopwords*) clear deactivated words.

(2)Number normalization

Since the numbers in the text should be matched the same, the numbers should be normalized during preprocessing. Convert different forms of numbers into Arabic numerals and keep them to two decimal places, as shown in Table 1.

Table 1. Example of number normalization

Number before conversion	Number after conversion
1/3	0.33
60%	0.60
25	52.00
28.472	28.47

(3)Extract feature vocabulary and negative vocabulary

Take the negative vocabulary as a benchmark, use regular matching to match the negative words in the text, and then extract the negative words of the sentence. LTP (*http://ltp.ai/index.html*) was used to extract the nouns, verbs, adjectives, adverbs, and syntactic analysis (the blue part in the figure) and the semantic dependency analysis of the sentence (the purple part in the figure), and the results are shown in Figure 2.

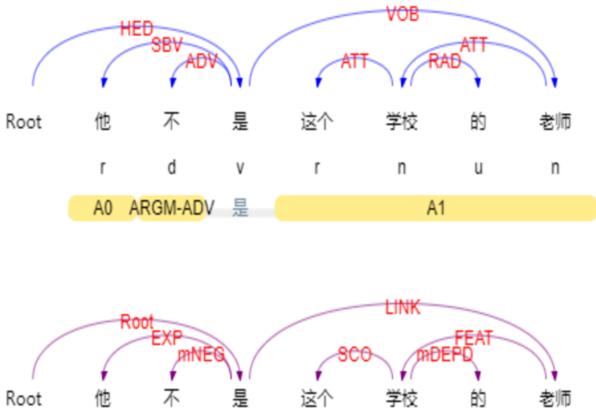

Figure 2. Sentence tagging, syntactic analysis and sentence semantic dependency analysis

3.2. **Semantic feature extraction.**

In the semantic feature extraction stage, Bi-LSTM is used to extract the contextual semantic relationship first, and then CNN is used to extract the local features of the text, and the text semantics are extracted at two levels to preserve as many semantic features of the text as possible.

3.2.1. *Semantic dependent feature extraction of context and context based on Bi-LSTM.* In order to extract the semantic dependencies of the context of the premise and the hypothesis, the BI-LSTM model, which is very suitable for modeling time series data, is used to train the premise and the hypothesis respectively, so as to better capture the long-distance dependencies. Using the Bi-LSTM model can better capture the dependence of longer distances [20–23]. Its internal network structure is shown as in Figure 3.

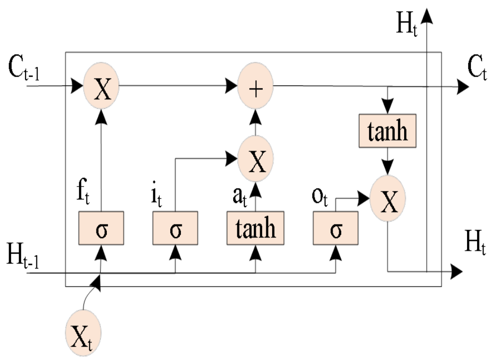

Figure 3. LSTM internal structure diagram

LSTM adds memory unit, input gate, output gate and forget gate on the basis of RNN.

The input of the network can be expressed as $[x_1, x_2, ..., x_T]$, the input gate i_t, output gate O_t, forget gate f_t and memory unit c_t in each LSTM network can be calculated by the following formula:

Let $X = [x_1, x_2, ..., x_T]$ is the input text, the realization of the internal structure of the LSTM neuron is as follows:

$$i_t = \sigma \left(W_{xi} x_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + b_i \right) \tag{1}$$

$$f_t = \sigma \left(W_{xf} x_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + b_f \right) \tag{2}$$

$$c_t = f_t c_{t-1} + i_t \tan h \left(W_{xc} x_t + W_{hc} h_{t-1} + b_c \right) \tag{3}$$

$$o_t = \sigma \left(W_{x0} x_t + W_{ho} h_{t-1} + W_{co} c_t + b_o \right) \tag{4}$$

$$h_t = o_t \tan h \left(c_t \right) \tag{5}$$

Among them: i_t, f_t, c_t, O_t, h_t are respectively the state of memory gate, hidden layer, forget gate, cell nucleus and output gate when the $t-th$ text is input. W_i, W_f, W_c, $W_o$are input gate weight, forget gate weight, cell state weight, and output gate weight, respectively. b_i, b_f, b_c, b_o are input gate bias, forget gate bias, cell state bias and output gate bias, respectively. W is the parameter of the model; b is the bias vector; σ s the Sigmoid function; *tanh* is the hyperbolic tangent function.

Bi-LSTM is a combination of forward LSTM and backward LSTM. The forward LSTM can memorize the above information, and the backward LSTM can memorize the following information, which promotes continuous sequence analysis. Bi-LSTM network structure diagram is shown as in Figure 4.

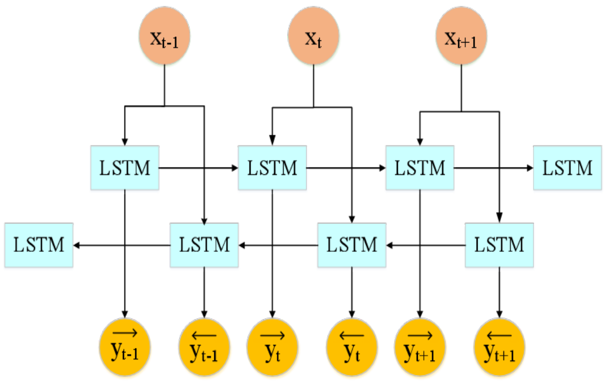

FIGURE 4. Bi-LSTM model diagram of bidirectional recurrent neural network

x_t in Figure 4 represents the input of the network at time t, and the LSTM in the box is the standard LSTM model. $\overrightarrow{y_t}$ is the output of the forward LSTM at time t, $\overrightarrow{y_t}$ is the output of the reverse LSTM at time t, that is to say, the output expression of Bi-LSTM at time t is defined as $y_t = [\overrightarrow{y_t}, \overleftarrow{y_t}]$, that is, the output at time t is directly spliced by the forward output and the reverse output.

This article uses the dual-tower model to train the premise sentence and the negative sentence separately. The length of a fixed sentence is 76 words (with 0 if it is less than 76), the hidden layer of Bi-LSTM has 256 neurons, so the output of Bi-LSTM is a matrix of (256, 76), the padding operation on this matrix is the (256, 256) matrix input into the CNN.

3.2.2. *CNN-based local feature extraction.* Bi-LSTM is better at learning the context-dependent features of the text, but considering feature vocabulary and negative vocabulary requires the model to be more sensitive to local features (such as learning the relationship between vocabulary). CNN has representation learning capabilities and can shift-invariant classification of input information according to its hierarchical structure. Therefore, CNN is added to extract the local features of the premise sentence and hypothesis sentence to facilitate subsequent semantic matching of the lexical semantic level. The CNN structure is shown in Figure 5.

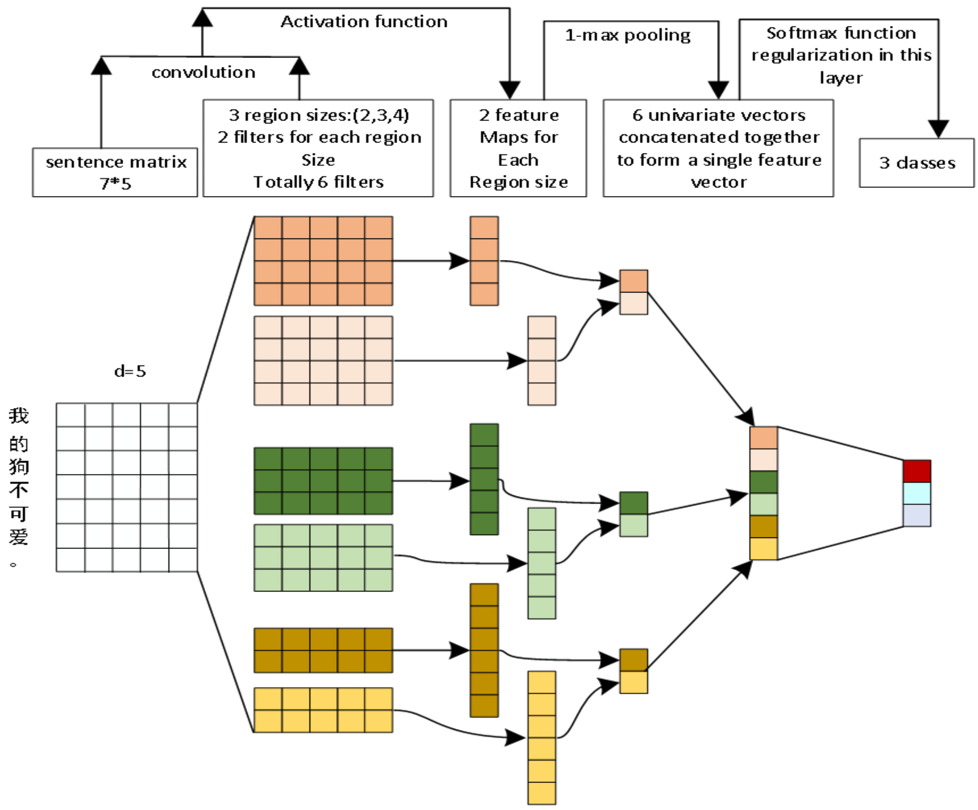

FIGURE 5. CNN structure

Since the Chinese text entailment recognition method that considers feature vocabulary and negative vocabulary adopts the dual-tower model, the CNN layer contains an input layer, a convolutional layer, a pooling layer, and a fully connected layer.

(1)Input layer

Taking the output vector of Bi-LSTM as the input of the CNN layer, the whole is the long-term dependence of the text extracted in the Bi-LSTM after the text is input into the CNN to extract the local information of the text.

(2)Convolutional layer

The convolution layer is composed of a series of convolution units (also called convolution kernel, filter), and the parameters of each convolution unit are obtained through the back propagation algorithm training. The convolutional layer can extract the local features of the input text because the input of each neuron is not fully connected to the previous layer, but connected to its local receptive field, so the local features are extracted. In view of the characteristics of the data set, this paper sets up three convolution kernels of different sizes, and determines the sliding step size as 2. In addition, in order to avoid the problem of different extraction times between the middle area and the edge area of the feature matrix, a padding operation is adopted (where Pad=1, that is, a circle of zeros is added to the periphery). This not only avoids less feature extraction in the edge area, but also makes the convolution output consistent with the input dimensions, and improves the maintainability and generalization of the program.

the input dimension of the convolutional layer is ω_{in}:

$$\omega_{in} = len_{in} \times \omega id_{in} \tag{6}$$

Where, len_{in} is the length of the input, ωid_{in} is the width of the input, then, the input dimension is calculated as follows:

$$\omega_{out} = len_{out} \times \omega id_{out} \tag{7}$$

$$len_{out} = \frac{len_{in} - Fliter + 2pad}{stride} + 1 \tag{8}$$

$$\omega id_{out} = \frac{\omega id_{in} - Fliter + 2pad}{stride} + 1 \tag{9}$$

From the above, $\omega_{out} = \omega_{in}$.

(3)Pooling layer

The feature dimension of the output of the convolutional layer is usually high, which will cause greater computational pressure. The way to solve this problem is to perform feature compression on the output of the convolutional layer through the pooling layer. This article chooses to use maximum pooling as the method of pooling. Maximum pooling means that when the filter slides to an area, the maximum value in the area is selected as the representation of the area. For the filter size selection, use a size of 2×2, and set the sliding step to 2. The original feature matrix is compressed to a quarter of the original size after the maximum pooling process, therefore, the pooling operation can effectively improve the calculation efficiency while maintaining the most salient features as much as possible.

(4)Fully connected layer

The fully connected layer receives a series of local features obtained after convolution and pooling operations, and reintegrates them into complete feature information through a weight matrix.

The input of CNN is a matrix of (256, 256). After a convolution, the output is a matrix of (128, 128). The output of the second convolution is a matrix of (65, 65). The output of the third convolution is (33, 33) matrix, and then MAX pools into a (4, 4) matrix, converts the (4, 4) matrix into a (1, 16) matrix and enters it into the preliminary recognition stage of text entailment. Because it is a dual-tower model, that is, the premise sentence and hypothesis sentence are converted into two (1, 16) matrices and input to the matching layer for text semantic matching.

3.3. **Preliminary recognition of text entailment.** In order to further improve the accuracy of text entailment recognition and reduce the dimensionality of the method, this paper adopts the dual-tower model to model. In the preliminary recognition of text entailment, semantic matching is performed on the premise sentence and hypothesis sentence, and then the entailment relationship between the two sentences is initially obtained. The preliminary recognition structure of text entailment is shown in Figure 6.

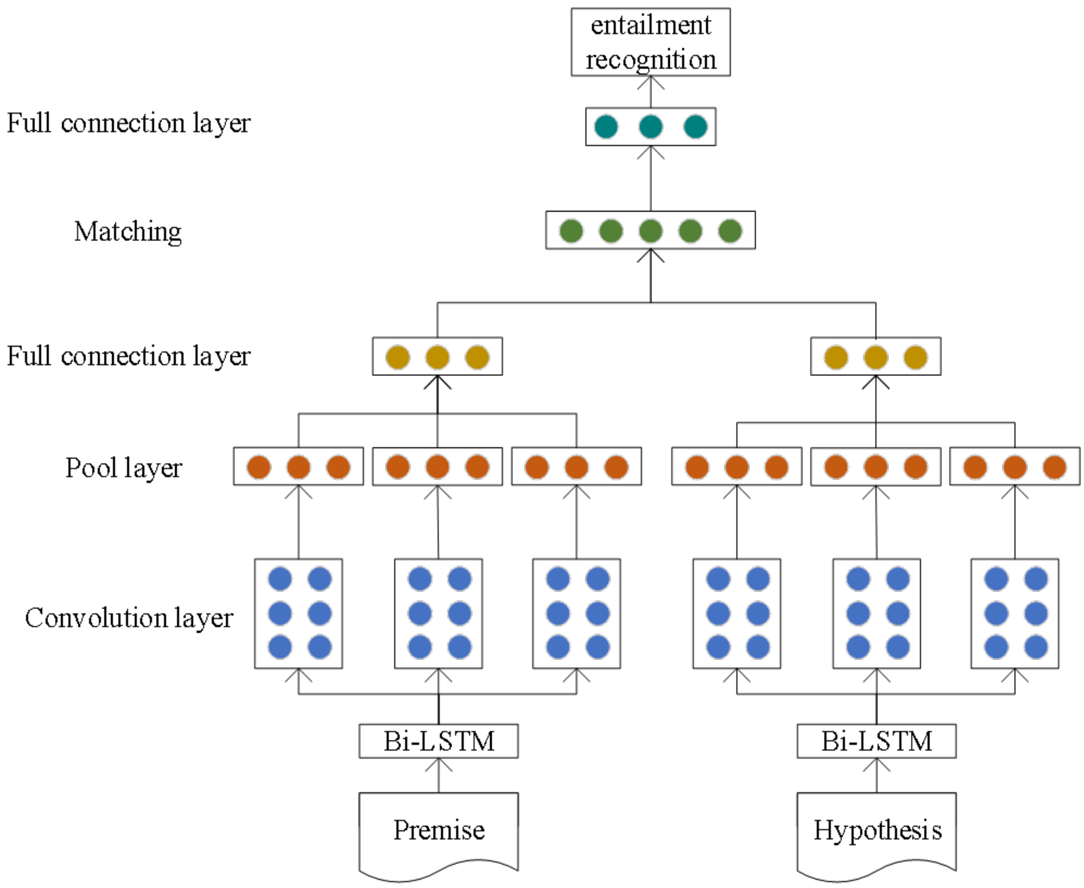

Figure 6. Structure diagram of preliminary recognition of text entailment

The single-tower structure separately extracts the semantic features of the premise sentence and hypothesis sentence, performs semantic matching in the matching layer, and then reduces to three dimensions through the fully connected layer, uses the softmax activation function to classify the text entailment recognition results, and finally outputs the text entailment recognition results.

3.4. **Result correction.** In order to improve the accuracy of the method for the recognition of text entailment containing negative words, the result correction stage is added to correct the recognition results. In the result revision stage, the preliminary recognition result is revised by adding the synonymous vocabulary, the antonymous vocabulary, the negative vocabulary, the number matching and the negative difference set.

Negative difference set refers to extracting the negative words of the premise sentence and hypothesis sentence respectively, and then using the *count*() function for statistics, and then calculating the parity of the negative difference by formula 10.

$$Parity = |P_p - P_h| \tag{10}$$

P_p indicates the number of negative words in the premise sentence, P_h indicates the number of negative words in the hypothesis sentence, *Parity* represents the parity of the difference in the number of negative words between the premise sentence and the hypothesis sentence. Add *Parity* as a parameter to the semantic feature extraction stage to update the weights of the semantic feature extraction parameters, and then pass the new weights into the text entailment preliminary recognition stage to update the preliminary recognition results, and iterate until the method reaches the optimal recognition Effect.

4. **Experiment and result analysis.** In order to verify the effectiveness and feasibility of the method proposed in this paper, 10,000 pieces of data containing negative words were randomly selected from the Liu Huanyong Chinese text entailment data set (*https://github.com/liuhuanyong/ChineseTextualInference*) and divided into training set and test set at 8:2 for experiments. This data set is comprehensively constructed based on the English text entailment data sets SNLI and MultiNLI through manual translation, machine translation, and manual sorting. Each piece of data contains two sentences (premise sentence and hypothesis sentence) and a label. The label is used to illustrate the relationship between the two sentences. The label has three categories, namely Entailment, Neutral and Contradiction.

The label distribution of the data is shown in Table 2:

Table 2. Data set label distribution

Label	Train	Test
Contradiction	2704	656
entailment	2642	684
neutral	2654	660
total	8000	2000

Table 2 shows that the label distribution of the data set is relatively uniform.

4.1. **Experimental environment.** The server operating environment used in the experiment is win10 system and python3.6; Use the jieba word segmentation tool to segment the text; use the Keras deep learning framework to complete the model development, and use the deep learning framework tensorflow1.14 as the running backend; the experimental specific parameter settings are shown in Table 3.

Table 3. Parameter settings

Parameter name	Parameter value
Batch	32
Epoch	50
Learning rate	0.001
Dropout	0.1

4.2. **Experimental evaluation index.** This article uses Accuracy, Precision, Recall, F1-Score as the evaluation indicators of the experiment. Among them, Accuracy indicates the probability of correct recognition for all recognized samples; precision represents the proportion of correctly identified samples among the samples that are identified as positive samples; recall represents the proportion of all positive samples that can be correctly identified; F1-Score is an indicator that neutralizes precision and recall. Accuracy, Precision, Recall, F1-Score formulas are as follows [24–28]:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} * 100 \tag{11}$$

$$\Pr ecision = \frac{TP}{TP + FP} * 100 \tag{12}$$

$$\operatorname{Re} call = \frac{TP}{TP + FN} * 100 \tag{13}$$

$$F1 - Score = \frac{2 * \Pr ecision * \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} * 100 \tag{14}$$

Among them, TP (True Positives) indicates that it is recognized as a positive example and is actually a positive example; FP (False Positives) means that it is recognized as a positive case, but it is actually a negative case; FN (false Negatives) means that it is recognized as a negative example, but it is actually a positive example; TN (True Negatives) indicates that it is recognized as a negative example and is actually a negative example.

4.3. **Analysis of experimental results.** In order to verify the effectiveness of the proposed method, 8 methods are selected for comparison with this method. Among them, BC1 represents single tower BI-LSTM +CNN, with correction layer, and feature words and negative words are extracted.BC2 represents bi-LSTM +CNN, with a correction layer, but no feature words and negative words are extracted.BC3 represents bi-LSTM +CNN with no correction layer, but feature words and negative words are extracted.BC4 represents bi-LSTM +CNN with no correction layer and no extracted feature words and negative words.BC5 stands for bi-LSTM +CNN, with correction layer, and feature words and negative words are extracted, that is, the method proposed in the text. The experimental results of various methods are shown in Table 4.3

TABLE 4. Comparison of experimental results (%)

contradiction	Bi-LSTM	Siamese+Bi-LSTM	ESIM	ABCNN	BC1	BC2	BC3	BC4	BC5
TP	53.21	38.24	40.01	36.35	61.02	60.87	61.25	53.37	61.93
FP	51.96	64.5	66.79	65.09	48.15	47.98	49.08	49.84	47.64
FN	46.79	61.76	59.99	63.65	38.98	39.13	38.75	46.63	38.07
TN	48.04	35.5	33.21	34.91	51.85	52.02	50.92	50.16	52.36
Precision	50.59	37.22	37.46	35.83	55.89	55.92	55.52	51.71	56.52
Recall	53.21	38.24	40.01	36.35	61.02	60.87	61.25	53.37	61.93
F1	51.87	37.72	38.69	36.09	58.34	58.29	58.24	52.53	59.10
ACC	50.63	36.87	36.61	35.63	56.44	56.45	56.09	51.77	57.15

entailment	Bi-LSTM	Siamese+Bi-LSTM	ESIM	ABCNN	BC1	BC2	BC3	BC4	BC5
TP	49.32	35.91	38.34	35.93	53.31	54.03	56.18	52.14	54.32
FP	49.67	63.37	59.74	64.34	42.15	43.05	44.03	47.27	40.22
FN	50.68	64.09	61.66	64.07	46.69	45.97	43.82	47.86	45.68
TN	50.33	36.63	40.26	35.66	57.85	56.95	55.97	52.73	59.78
Precision	49.82	36.17	39.09	35.83	55.85	55.66	56.06	52.45	57.46
Recall	49.32	35.91	38.34	35.93	53.31	54.03	56.18	52.14	54.32
F1	49.57	36.04	38.71	35.88	54.55	54.83	56.12	52.29	55.84
ACC	49.83	36.27	39.30	35.80	55.58	55.49	56.08	52.44	57.05

neutral	Bi-LSTM	Siames+Bi-LSTM	ESIM	ABCNN	BC1	BC2	BC3	BC4	BC5
TP	48.4	34.51	36.37	35.43	53.75	52.71	50.93	50.7	55.71
FP	47.44	63.47	58.75	62.96	41.62	41.36	38.53	46.68	40.18
FN	51.6	65.49	63.63	64.57	46.25	47.29	49.07	49.3	44.29
TN	52.56	36.53	41.25	37.04	58.38	58.64	61.47	53.32	59.82
Precision	50.50	35.22	38.24	36.01	56.36	56.03	56.93	52.06	58.10
Recall	48.4	34.51	36.37	35.43	53.75	52.71	50.93	50.7	55.71
F1	49.43	34.86	37.28	35.72	55.02	54.32	53.76	51.37	56.88
ACC	50.48	35.52	38.81	36.24	56.07	55.68	56.20	52.01	57.77

It can be seen from Table 5 that the average accuracy rate of the Chinese text entailment recognition method (BC5), which considers characteristic vocabulary and negative vocabulary, is higher than that of other mainstream methods. Especially in the recognition of neutral type endoment relations, the accuracy rate reached 57.77%, the recall rate reached 55.71%, and the accuracy rate reached 58.10%. The effect on the recognition of contradiction entailment relations is also significant, with an accuracy rate of 57.15%, a recall rate of 61.93%, and an accuracy rate of 56.52%. The second is BC3 (dual-tower Bi-LSTM+CNN, no correction layer, but feature vocabulary and negative vocabulary are extracted). The effect is better, The accuracy rates of contradiction type, entailment type, and neutral type are 56.09%, 56.08% and 56.20%, the recall rates are 61.25%, 56.18% and 50.09%, and the precision rates are 55.52%, 52.45% and 56.93%. It can be explained that enriching semantics from the perspective of characteristic vocabulary and negative vocabulary can significantly improve the recognition effect of text entailment.

The confusion matrix of the experimental results of various methods is shown in Figure 7. Among them, (a)-(i) respectively represent Bi-LSTM method, Siamese +Bi-LSTM method, ESIM method, ABCN method, BC1 method, BC2 method, BC3 method, BC4 method and BC5 method.

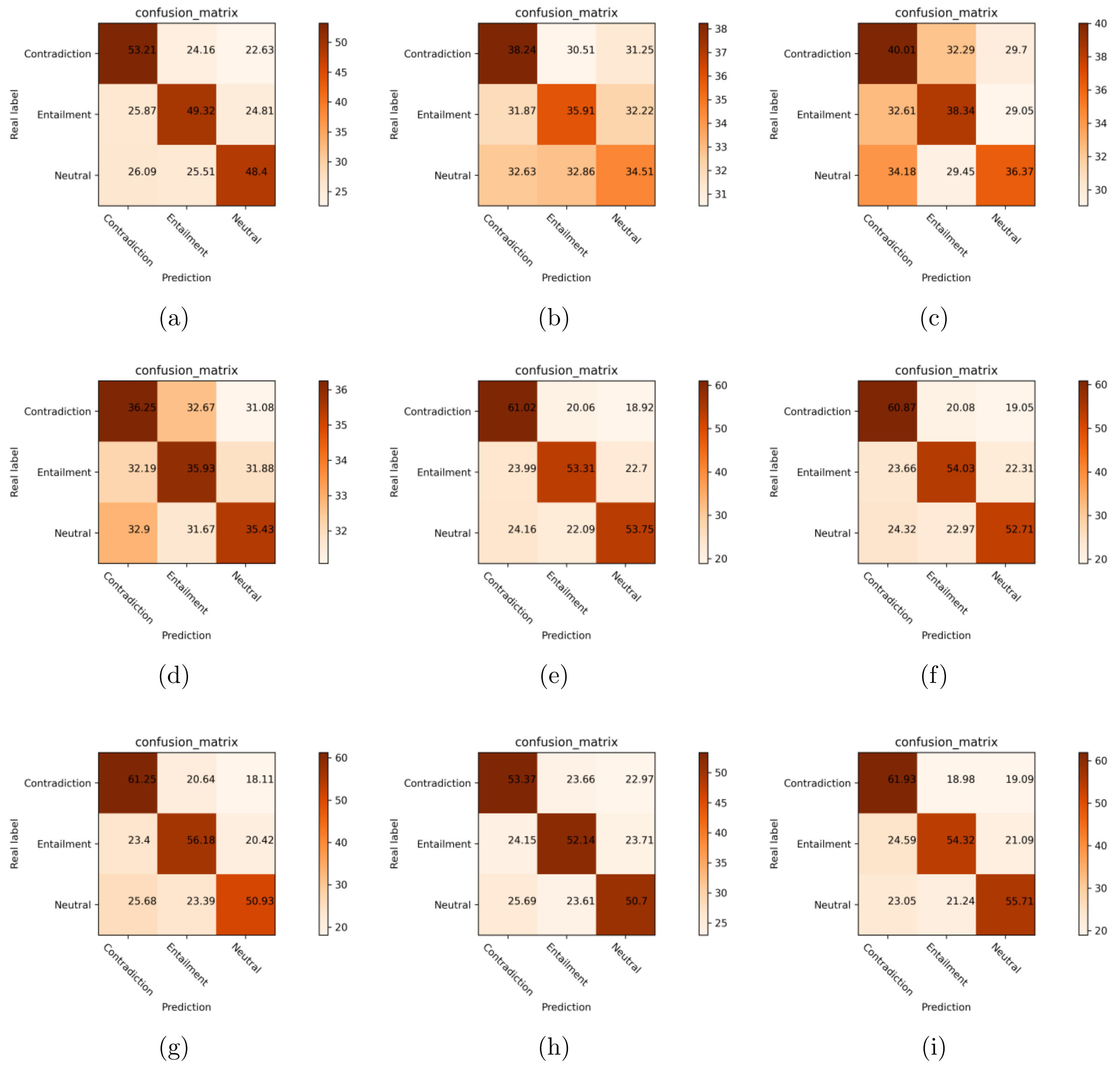

FIGURE 7. Confusion matrix of experimental results

It can be seen from Figure 7 that these nine methods all have a higher recognition accuracy for contradiction-type Chinese text entailment with negative words, followed by neutral relations, but the overall accuracy needs to be improved. The overall recognition accuracy of Siamese +Bi-LSTM method, ESIM method and ABCNN method is lower than Bi-LSTM method, BC1 method, BC2 method, BC3 method, BC4 method and BC5 method. It shows that the mainstream method of Bi-LSTM+CNN has a better effect on the recognition of Chinese text entailment. The Siamese +Bi-LSTM method has the worst effect on the recognition of neutral text entailment, while the contrast recognition effect of text entailment is better; The ESIM method easily recognizes the text as a contradiction type; Among the samples with incorrect recognition by the ABCNN method, the probability of identifying neutral as contradiction is the largest, reaching 32.90%; The BC5 method has the best performance in the recognition of contradiction, entailment and neutral text relations, among the samples with incorrect recognition, the probability of recognizing the entailment type as contradiction is the largest, reaching 24.59%.The comprehensive correct rate of different methods is shown in Figure 8.

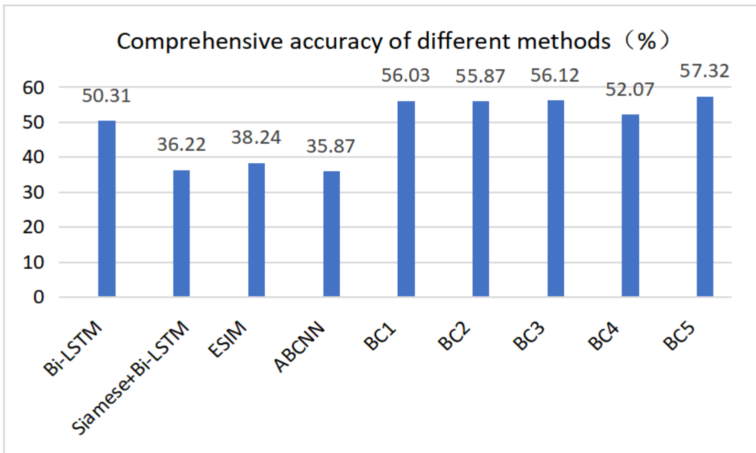

FIGURE 8. Comprehensive accuracy of different methods

It can be seen from Figure 8 that the experimental accuracy of the dual-tower model is about 2% higher than that of the single-tower model (comparison between the BC1 method and the BC5 method); adding CNN to extract local features of text can increase the accuracy of text entailment by about 2% (comparison between Bi-LSTM method and BC4 method); adding a correction layer can increase the accuracy of text entailment by about 1% (comparison between BC3 method and BC5 method); extracting characteristic vocabulary and negative vocabulary can increase the accuracy of text entailment by about 2% (comparison between BC2 method and BC5 method);Extracting characteristic vocabulary and negative vocabulary and adding a correction layer can increase the accuracy of text entailment by about 5% (comparison between BC4 method and BC5 method); the Siamese +Bi-LSTM method, ESIM method and ABCNN method are not effective in the recognition of Chinese text entailment. Experiments prove that our method is effective, and the comprehensive accuracy rate can reach 57.32%, which is significantly better than other methods.

4.4. **Discussion.** By comparing the recognition effect of Chinese text entailment with negative words on different methods, the Chinese text entailment recognition method that considers characteristic vocabulary and negative vocabulary proposed in this paper has a better effect. Compared with the method that does not add the correction layer and does not extract the feature vocabulary and negative vocabulary of the text, the accuracy rate is improved by 5%, this method is especially effective in recognizing contradiction-type Chinese text entailment, with an accuracy rate of 61.93%.

A retrospective analysis of 500 sentence pairs with recognition errors found that the samples with recognition errors can be mainly divided into three types. They are calculated text (76/500), common sense text (153/500) and complex text (162/500).

Computational text means that when mathematical calculations are involved in the text, this method is often not able to accurately calculate. For example: Premise sentence: 100 people took the test and 31 students failed the test. Hypothesis sentence: About two-thirds of the students passed this exam. The Chinese text entailment recognition method that considers characteristic vocabulary and negative vocabulary will recognize these two sentences as a contradiction relationship (actually an entailment relationship).

Common sense text refers to when the text involves the introduction of external knowledge, this method is often not able to judge the entailment relationship of the text well. For example: Premise sentence: each prisoner can be locked up for only $25,000 per year, and this does not include any medical or dental health they receive. Hypothesis sentence: prisoners spend a lot of money for society, and this money could have been better spent on health care. The Chinese text entailment recognition method that considers characteristic vocabulary and negative vocabulary will recognize these two sentences as an entailment relationship (actually a neutral relationship).

Complex text means that when the grammar of the text is complex or the sentence is long, this method is often not able to judge the entailment relationship of the text well. For example: Premise sentence: as his partner, he will become misunderstood. There is no doubt that in the Disney version, Beowulf and the warrior princess will inevitably be written into the script. Glen Beowulf and Grendel become friends(It is not easy). Hypothesis sentence: Disney will undoubtedly portray Del and Grendel's mother's friends. The Chinese text entailment recognition method that considers characteristic vocabulary and negative vocabulary will recognize these two sentences as a neutral relationship (actually an entailment relationship).

Based on the above analysis, when doing text entailment, you can consider digital calculation, external knowledge introduction, and more detailed text semantic mining.

5. **Conclusions.** Due to the low recognition accuracy of Chinese text entailment containing negative words, this paper proposes a Chinese text entailment recognition method that considers characteristic vocabulary and negative vocabulary. The innovation of this method is that the influence of characteristic words (nouns, verbs, adjectives, adverbs) and negative words on the semantics of the text is considered in the text preprocessing stage, as well as the comprehensive application of the dual-tower model and correction. Experimental results show that because this method enriches the text semantics, it has higher accuracy in recognizing text entailment that contains negative words than other methods. This method adds the result correction stage to correct the entailment recognition results, which also improves the accuracy of entailment recognition. Experiments have proved the effectiveness and feasibility of the method proposed in this paper. However, there is still room for improvement in this method. When conducting a retrospective analysis of the experiment, it was found that the recognition effect of this method is not good on computational text, common sense text, and complex text. In future research, more detailed research can be considered in digital computing, the introduction of external knowledge, and semantic mining. In addition, there are currently few Chinese text entailment data sets, and there are almost no Chinese text entailment data sets specific

to a certain field, which makes current research not comprehensive enough. Therefore, it is also a meaningful work to establish a Chinese text data set in a specific field.

6. **Acknowledgment.** This article is supported by the National Key R&D Program Project "Multi-value Chain Collaborative Data Space Management Engine and Management System Architecture Design and Verification" (No.2020YFB1707804), and the 2020 Jilin City Science and Technology Development Plan Project "Jilin City Tourism Online Comment Text Emotion Classification Research"(No.20200104108).

REFERENCES

[1] J. Peng, D.Q. Yang, S.W. Tang, A new similarity computing method based on concept similarity in Chinese text processing, *Science in China Series F: Information Sciences*, vol.51, no.9, pp.1215-1230, 2008.
[2] W. Guo, Q. Zeng, H. Duan, Process-Extraction-Based Text Similarity Measure for Emergency Response Plans, *Expert Systems with Applications*, vol.183, 115301, 2021.
[3] Inan E. Inan, SimiT: A Text Similarity Method Using Lexicon and Dependency Representations, *New Generation Computing*, vol.38, no.3, pp.509-530,2020.
[4] V. Jijkoun, M.D. Rijke, Recognizing textual entailment using lexical similarity, *Proceedings of the PASCAL Challenges Workshop on Recognising Textual Entailment*, pp.73-76, 2005.
[5] D. Lin, Extracting collocations from text corpora, *First workshop on computational terminology*, pp.57-63, 1998.
[6] G.A. Miller, WordNet: a lexical database for English, *Communications of the ACM*, vol.38, no.11, pp.39-41, 1995.
[7] M. Kouylekov, M. Negri, An open-source package for recognizing textual entailment, *Proceedings of the ACL 2010 System Demonstrations*, pp.42-47, 2010.
[8] M. Heilman, N.A. Smith, Tree edit models for recognizing textual entailments, *paraphrases, and answers to questions, Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp.1011-1019, 2010.
[9] J.R. Hobbs, M.E. Stickel, D.E. Appelt, Interpretation as abduction, *Artificial intelligence*, vol.63, no.2, pp.69-142, 1993.
[10] Bos, Johan, Is there a place for logic in recognizing textual entailment, *Linguistic Issues in Language Technology*, vol. 9, pp. 27-44, 2014.
[11] D. Moldovan, C. Clark, S. Harabagiu, Cogex: A logic prover for question answering, *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, pp.166-172, 2003.
[12] R. Raina, A.Y. Ng, C.D. Manning, Robust textual inference via learning and abductive reasoning, *Proceedings of the 20th national conference on Artificial intelligence*, vol. 3, 10991105, 2005.
[13] Y. Kim, Convolutional Neural Networks for Sentence Classification, arXiv:1408.5882
[14] N. Kalchbrenner, E. Grefenstette, P. Blunsom, A convolutional neural network for modelling sentences, arXiv:1404.2188
[15] L. Mou, R. Men, G.Li, Natural language inference by tree-based convolution and heuristic matching, arXiv:1512.08422
[16] S.R. Bowman, G. Angeli, C. Potts, A large annotated corpus for learning natural language inference, arXiv:1508.05326
[17] S.R. Bowman, G. Angeli, C. Potts, A fast unified model for parsing and sentence understanding, arXiv:1603.06021
[18] J. Devlin, M.W. Chang, K. Lee, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv:1810.04805
[19] Y. Liu, C.J. Sun, L. Lin, Computing semantic text similarity using rich features, *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation*, pp.44-52, 2015.
[20] Y.Q. Chen, B. Zhou, M.M. Zhang, C.M. Chen, Using IoT Technology for Computer-Integrated Systems in the Semiconductor Industry, *Applied Soft Computing*, vol. 89, 106065, 2020.
[21] S. Kumar, A. Damaraju, A. Kumar, LSTM Network for Transportation Mode Detection, *Journal of Internet Technology*, vol.22, no.4, pp.891-902, 2021.
[22] Wu, M.T. Jimmy, A graph-based CNN-LSTM stock price prediction algorithm with leading indicators, *Multimedia Systems*, pp.1-20, 2021.

[23] Wu, M.T. Jimmy, Convert index trading to option strategies via LSTM architecture, *Neural Computing and Applications*, pp.1-18, 2020.

[24] T.Y. Han, X. Chen, Q.X. Liu, Research on short-term load forecasting of distribution network based on long and short-term memory network under peak and valley electricity price, *Journal of Northeast Electric Power University*, vol.40, no.04, pp.19-28, 2020.

[25] J. Gao, Y.D. Wang, R.D. Han. Optimal Stop Time in Data Information Proofreading Decision-Making, *Journal of Northeast Electric Power University*, vol.40, no.06, pp.105-108, 2020.

[26] W. Tang, S. Wu, R. Li, X. Jin, D. Yang, Z.Y. Liu, W.P. Hong, Optimization of District Cooling Pipeline Network Based on Genetic Algorithm, *Journal of Northeast Electric Power University*, vol.40, no.06, pp.86-91, 2020.

[27] Y. Zhang, R. Zhang, Mathematical model and artificial intelligence of multidimensional database and spatial multidimensional data, *Journal of Northeast Electric Power University*, vol.40, no.05, pp.84-92, 2020.

[28] P. Chen, M. Xia, A Java Contractual Programming Language Model Using AspectJ Technology, *Journal of Northeast Electric Power University*, vol.31, no.03, pp.52-55, 2011.