Hindawi Scientific Programming Volume 2021, Article ID 4176059, 8 pages https://doi.org/10.1155/2021/4176059



Research Article

The Research of Convolutional Neural Network Based on Integrated Classification in Question Classification

Lihua Zhen 101 and Xiaoqi Sun2

¹School of Economics and Management, Nantong University, Nantong 226000, China ²School of Educational Science, Nantong University, Nantong 226000, China

Correspondence should be addressed to Lihua Zhen; zlh1988@ntu.edu.cn

Received 14 September 2021; Revised 30 September 2021; Accepted 1 October 2021; Published 12 October 2021

Academic Editor: Tongguang Ni

Copyright © 2021 Lihua Zhen and Xiaoqi Sun. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a new generation of search engine, automatic question answering system (QAS) is becoming more and more important and has become one of the hotspots of computer application research and natural language processing (NLP). However, as an indispensable part of the QAS, the role of question classification is an understood thing in the system. In view of this, to further make the performance of question classification much better, both the feature extraction and the classification model were explored. On the study of existing CNN research, an improved CNN model based on Bagging integrated classification ("W2V + B-CNN" for short) is proposed and applied to question classification. Firstly, we combine the characteristics of short texts, use the Word2Vec tool to map the features of the words to a certain dimension, and organize the question sentences into the form of a two-dimensional matrix similar to the image. Then, the trained word vectors are used as the input of the CNN for feature extraction. Finally, the Bagging integrated classification algorithm is used to replace the Softmax classification of the traditional CNN for classification. In other words, the good of W2V + B-CNN model is that it can make use of the advantages of CNN and Bagging integrated classification at the same time. Overall, the new model can not only use the powerful feature extraction capabilities of CNN to extract the potential features of natural language questions but also use the good data classification capabilities of the integrated classification algorithm for feature classification at the same time, which can help improve the accuracy of the W2V + B-CNN in the application of question classification. The comparative experiment results prove that the effect of the W2V + B-CNN is significantly better than that of the CNN and other classification algorithms in question classification.

1. Introduction

In the Internet age, information has exploded. In the face of massive fragmented information, people's desire to quickly obtain accurate and concise information has become more and more urgent, and the QAS has emerged at the historic moment. Unlike traditional search engines, QAS is a highlevel form of information retrieval, which has become a hot focus in the field of current natural language. It allows users to describe the problems in natural language and can find or infer the answers to the users' questions from massive heterogeneous data and then submit them to the users. For example, for the question "What color is the skin of Chinese people," the system will give the answer directly as "yellow."

It greatly improves the users' query efficiency and better meets the users' needs. The QAS generally includes three main parts: question analysis, information retrieval, and answer extraction [1], and each part cooperates with each other to efficiently obtain the target information required by the user. If question analysis is the cornerstone of QAS, then question classification is not only one of the key parts of question analysis but also an indispensable module of the QAS. It not only helps to optimize the performance of the system, such as reducing the search space of candidate answers and the time to locate the correct position of answers, but also helps to formulate answer extraction strategies. It can be seen that the result of question classification can provide useful guidance information for other modules

of the system, and its accuracy will have a direct influence on the quality of the whole QAS. Therefore, it can be said that the research on question classification has important practical significance and has a positive influence on improving the quality and performance of QAS.

Question classification is to determine the type of question associated with the characteristics of the answer or the semantic information of the question under a certain classification standard. Currently, with the persistent maturity and improvement of its research, it has also attracted much attention in the field of NLP. The essence of question classification is short text classification. The current research on it is generally based on the idea of text classification and combined with the characteristics of question classification itself. But different from ordinary text classification, question classification contains unique word feature information. How to fully mine this information is the key to question classification. There are two types of traditional question classification methods: rule-based methods and statistical machine learning methods. Early rule-based methods mainly used artificial analysis of the syntactic structure to extract rules and then judge the question type [2, 3]. This kind of method has many advantages. For example, it is relatively easy to implement and does not need a lot of training data, so the classification speed is fast. But the disadvantage is that these methods rely more on experts and are subjective. Moreover, the classification decision of experts is very easy to be affected by the classification system, which makes it less flexible. Subsequently, the methods based on statistical learning showed a good classification effect, which have the advantages of strong versatility, easy transplantation, and expansion [4, 5]. The machine learning models based on statistical methods commonly used in question classification are Bayes [6], SVM [7, 8], KNN [9, 10], ME [11], and so on. However, the disadvantage of the statistical learning method is that its classification accuracy is still easily affected by the accuracy of syntactic analysis.

Recently, with the wide application of deep learning technology in the field of NLP, its characteristics of not relying on complex feature engineering and fully mining the feature information of natural language have attracted the interest of a large number of scholars. They began to use deep learning methods to classify questions. Therefore, question classification methods based on deep learning technology came into being. Different from the previous classification methods, the DNN model has significant advantages in question representation and feature extraction. The multilayer network structure can abstract the original question sentence into a high-level vector representation so that the dimension of the feature vector becomes higher and the classification accuracy is greatly improved. Meanwhile, with the wide study of word vector technology and deep neural network technology, people have more new ideas in NLP tasks. Among the deep neural networks applied to NLP tasks, the most commonly used ones are RNN and CNN. Among them, CNN is a typical spatial deep neural network. It has significant advantages in feature extraction and is good at feature self-extraction, which can reduce the difficulty of feature extraction to a

certain extent in question classification and improve the accuracy of classification. Based on this, it has inspired researchers to use CNN as one of the commonly used architectures in the application of deep learning in question classification. Many methods based on CNN have been proposed and a large number of research results have emerged [12, 13]. Therefore, there is still a lot of research space for question classification with deep learning, which is worthy of further research.

2. Related Work

2.1. Question Classification. The purpose of question classification is to classify questions into corresponding semantic classes according to the type of answer. As a key link of the question answering system, it has an important guiding role for subsequent answer extraction modules, and it also has a very important significance for the QAS [14]. The so-called question classification means that, under a certain classification system, for questions that are not marked with a class, the system automatically classifies the relevant class of the question according to the content of the question. This correspondence can be abstracted as a mapping process in a mathematical sense, which can be represented by the following mapping function [15]:

$$R: X \longrightarrow \{C_1, C_2, \dots, C_N\},\tag{1}$$

where X represents a set of question samples and $\{C_1, C_2, \ldots, C_N\}$ represents a set of question classes. R is responsible for mapping the question $x \in X$ of the unknown class to a class C_i according to a certain rule or a certain classification algorithm.

However, the current methods based on question classification research usually draw on ideas of text classification. The difference between them is that common words such as "what" and "is" are often overlooked in text classification, but these words that may be used as stop words are often very important in question classification. This is also one of the characteristics of question classification. What the two have in common is that they both analyze the information contained in the text and combine the characteristics of question classification to classify the questions into their categories. Based on the ideas of text classification, the question classification process can be described as shown in Figure 1, which specifically includes the division of training set and test set, preprocessing, feature extraction, classifier training, classification prediction, and others [16]. For Chinese text, data preprocessing includes Chinese word segmentation, part-of-speech tagging, and stop word removal. The advantage of feature extraction is that it helps to reduce the complexity and improve the accuracy of question classification by extracting better feature information from the original samples. Common methods include TF-IDF calculation, n-gram, Word2Vec, and LDA. Since question classification is generally a multiclass model, machine learning generally uses methods like Bayes, KNN, and SVM for classification. The flowchart of the question classification process is shown in Figure 1.

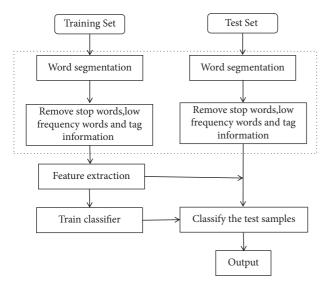


FIGURE 1: The flowchart of the question classification process.

2.2. Question Classification Based on Deep Learning. The concept of deep learning was first proposed by Hinton and Salakhutdinov [17] and others in 2006. It stimulates the mechanism of the human brain and can extract features layer by layer from the original data. To a large extent, it is helpful for solving the problems of time-consuming, laborious, and poor effectiveness that often occur in traditional machine learning methods which need to customize feature extraction rules by themselves. In recent years, with continuous development of deep learning technology, its application value in the fields of image processing [18], pattern recognition [19], and NLP is self-evident. For question classification tasks, deep learning can be used to actively analyze and learn the syntactic and semantic features implicit in the questions. This helps us to analyze the semantic feature structure of the question at a deeper level and perform feature extraction, so as to make the question classification more accurate. Unlike traditional machine learning methods, deep learning does not require manual extraction of the characteristics of question sentences, which greatly reduces labor and time costs. It can automatically obtain basic features, then combine these basic features into complex features, and finally train the model to judge the semantic relationship between question features and question classes. Therefore, the method based on deep learning has faster data processing capabilities and strong adaptive deep learning capabilities. Meanwhile, its fault tolerance and noise resistance are relatively high, which is very suitable for question classification.

Currently, the most representative models in deep learning for question classification are CNN model, long short-term memory network (LSTM) model, and Bi-LSTM model. What is more, a large number of relevant research results have also emerged. Kim et al. introduced a sentence classification method with a convolutional neural network based on word vectors in 2014. They used CNN to classify English questions by turning question sentences into word vectors [20]. After that, Zhang et al. further improved Kim's

model and proposed a novel low-complexity model termed CNN-BiGRU [21]. They introduced the bidirectional gated recurrent unit (BiGRU) into the traditional CNN model to naturally learn the question sentence and achieve classification. It improves the classification accuracy of the CNN model on a variety of English classification data sets. Kalchbrenner et al. [12] constructed a Dynamic Convolutional Neural Network (DCNN). In the network, they used a global k-max pooling operation to solve the problem of inconsistent question length, used the DCNN network to simulate the semantic information of the question, and achieved a better question classification effect. Le and Zuidema studied the problems of incorrect modification of syntactic components and incorrect conversion of syntactic tree in the question classification of Recurrent Neural Network (RNN). They used the syntactic forest as the input of the CNN network, proposed the Forest Convolutional Network (FCN), and achieved good results in the question classification task [22]. Nan et al. constructed a neural network model based on LSTM to complete the joint modeling of the description subject and the description text and achieved a good classification effect [23].

Inspired by the above research, we started research from two aspects of feature extraction and classification model. On the basis of CNN, we further integrated the advantages of the Bagging classification algorithm, proposed an integrated convolutional neural network model, and applied it to the problem classification. The specific research content is as follows:

- (1) Firstly, we use Word2Vec and CNN to complete the feature representation and feature extraction of question sentences. Since the question contains fewer words, the traditional vector space model may cause problems, such as the feature latitude that is too high and or feature vector data that is sparse. Moreover, since the correlation between words and the position information of the words in the document is not considered, these factors will affect the accuracy of question classification. Therefore, we combine the characteristics of short texts, use the Word2Vec tool to map the features of the words to a certain dimension, organize the question sentences into the form of a two-dimensional matrix similar to the image, and then set the matrix as the input of the CNN. Then, we design the CNN model and complete the feature extraction of the input data through operations such as convolution and pooling.
- (2) Based on the study of the CNN model, the Bagging algorithm is used to construct the classification layer. Aiming at the problem of weak generalization ability caused by the Softmax classifier used by most convolutional neural networks at present, we propose a B-CNN integrated model combined with the Bagging algorithm to improve the accuracy and generalization of question classification.
- (3) Finally, to verify the effectiveness and feasibility of the new model in question classification, several

experiments were conducted in the Chinese question set provided by the Information Retrieval Laboratory of Harbin Institute of Technology. The results prove that the effect of the W2V + B-CNN model on question classification is significantly better than that of the CNN and other classification algorithms.

3. Question Classification Model of CNN Based on Bagging Integrated Classification

To optimize the accuracy of the CNN model in processing multiclassification of questions, the paper further combines the advantages of the Bagging algorithm on the basis of the CNN and proposes an integrated convolutional neural network model referred to as W2V + B-CNN. Meanwhile, we apply it to question classification. The basic principle of W2V + B-CNN is to express the question as a word sequence with words as the unit and map it to a multidimensional vector to construct a set of word vectors. Furthermore, the characteristic information in the question is extracted through the integrated neural network, so as to realize the question classification. The structure diagram of the W2V + B-CNN model used in question classification in this paper is shown in Figure 2.

In Figure 2, we can see that the model is mainly composed of three layers, namely, the word vector matrix input layer, the convolutional feature extraction layer (including convolution layer and maximum pooling layer), and the integrated classification layer (including Dropout and Bagging integrated classification output). The word vector matrix input layer uses the Word2Vec tool to train the input sentence, converts the words into word vectors, and then splices them into a text word vector matrix. The convolution layer uses the convolution kernel to perform convolution operation on the input feature vector to extract features. The pooling layer performs sampling processing on the features extracted by the upper layer and retains important features in the form of filtering. In the integrated classification layer, the pooled and spliced feature vectors should be dropout processed during model training. Finally, the Bagging integrated classifier is used to complete the mapping of the feature vector to the category, so as to obtain the final classification result.

3.1. Word Vector Matrix Input Layer. Word vector can also be called word embedding technology, which can map words containing rich semantic information to abstract high-dimensional vector space. It is a method of continuous digital vectorization of words using a shallow neural network [24, 25]. The advantage of word vector technology is that it is helpful in solving the problem of data sparsity in traditional question classification methods. In the research of a large number of word vector learning methods, Mikolov and others from Google open-sourced a tool for generating word vectors called Word2Vecor in 2013, which includes two models of CBOW and Skip-gram [26, 27]. The structure of the two models is shown in Figure 3. The modeling idea of CBOW is to use the words in the window to predict the

central word, while the modeling idea of Skip-gram is to use the central word to predict the surrounding words. In our study, the Skip-gram model will be chosen to train word vectors.

Assume that, after training, each question S can be represented by a word vector matrix, that is, $S = \{s_1, s_2, \dots, s_n\}$, where s_i is the ith word of question S and n represents the number of words contained in S. Each word s_i can be represented by a word vector, that is, $\{w_1, w_2, \dots, w_k\}$, where w_i is the weight of the ith dimension in the word vector and k represents the dimension of the word vector. Two models of CBOW and Skip-gram are shown in Figure 3.

3.2. Convolutional Layer. The convolutional layer is the core component of the CNN, and the core point is to capture the local correlation. Specifically for question classification tasks, the convolution kernel can extract key information similar to N-grams in the sentence. The convolution operation in the text is somewhat different from the convolution operation in the image. Because a word vector is a whole in the text, it makes sense to perform a convolution operation on the whole word vector. Therefore, assuming that the dimension of the word vector is k, then the width of the convolution kernel should also be k. When convolution operation is performed on the sentence $S = \{s_1, s_2, \dots, s_n\}$, the convolution kernel W can be expressed as $W \in \mathbb{R}^{hk}$, and h is the height of the convolution kernel. Then, every time you slide a word vector matrix with length h and width n, you will get an eigenvalue. The calculation formula of the eigenvalue is

$$d_i = f\left(S_{i:\ i+h-1} \cdot W + b\right),\tag{2}$$

where b is a bias term, $S_{i: i+h-1}$ is a word sequence of length h, namely, $(s_i, s_{i+1}, \ldots, s_{i+h-1})$, and f is an activation function. Commonly used activation functions include tanh, ReLU, and sigmoid. Increasing the activation function can make the model introduce nonlinear factors, so as to better fit the data. And the ReLU activation function will be applied in our study. After the entire question sentence is subjected to convolution operation, several feature vectors representing the sentence can be obtained:

$$D = (d_1, d_2, \dots, d_{n-h+1}). \tag{3}$$

Because convolution kernels of different sizes can extract the features of questions from different angles, the feature information extracted by the convolution layer is affected by the size of the convolution kernel.

3.3. Pooling Layer. The pooling layer can reduce the dimensionality of features by downsampling the output vector of the convolutional layer. On the one hand, it can speed up the calculation, and on the other hand, it can effectively prevent the problem of simulation overfitting. In this paper, the maximum pooling method is used to process the feature vector D obtained by the convolutional layer, so as to select the most representative feature. The formula is

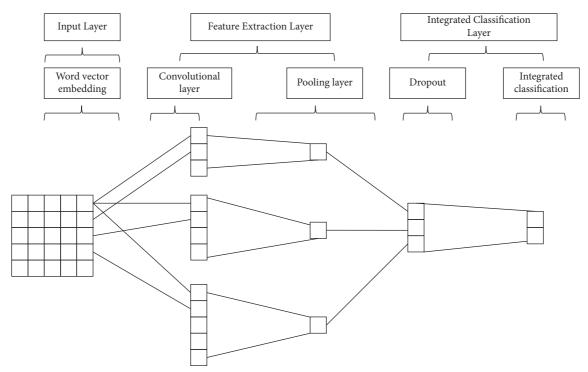


FIGURE 2: The structure diagram of the W2V + B-CNN model.

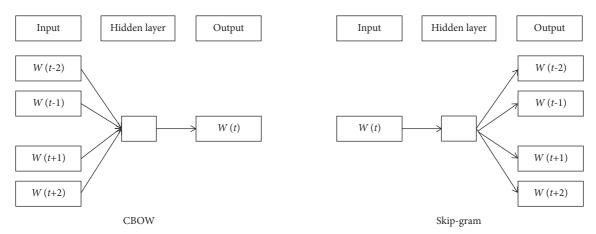


FIGURE 3: Two models of CBOW and Skip-gram.

$$D_{\max} = \max(D). \tag{4}$$

3.4. Integrated Classification Layer

3.4.1. Dropout. To avoid the problem of overfitting in the process of training, the Dropout operation is usually used to prohibit some hidden nodes from participating in forward propagation. These neurons will not participate in the update process so that the update of the weight will not rely on the role of the fixed nodes.

3.4.2. Bagging Integrated Algorithm. To further optimize the classification ability of the CNN, the Bagging integrated

algorithm with better classification performance will be utilized to replace the Softmax classification function in the CNN. In the integrated classification layer, we will first apply the trained features of the convolutional layer and pooling layer as a new feature set, then input them into the Bagging integrated learning classifier for training, and finally output the classification results according to the voting method. In this way, CNN can be used to extract the potential features of the data set, and the integrated learning can be used for feature classification, which can help improve the accuracy of multiclassification tasks.

Suppose that, after the Dropout operation, the data set input to the integrated classifier is $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\}$. Define $Y = \{-1, +1\}$ as the set of classification labels, T represents the number of

base classifiers, and L represents the base classifiers. The hypothetical function after integration with the Bagging method is

$$H(x) = \arg\max_{y \in Y} \sum_{i=1}^{T} sign\left(\sum_{i=1}^{T} h_i(x) = y\right), \tag{5}$$

where h_i (i = 1, 2, ..., T) represents the bootstrap distribution and its formula is

$$h_i = \frac{L(D, D_{bs})}{D_{bs}}. (6)$$

4. Experimental Results and Analysis

4.1. Experimental Data. To verify the performance of the W2V B-CNN model, we designed several experiments on the Chinese question set provided by the Information Retrieval Laboratory of Harbin Institute of Technology. It is a relatively classic Chinese question classification data set with good universality and can better prove the performance of the algorithms. The classification system of the question set is divided into 7 classes, including description (DES), human (HUM), location (LOC), number (NUM), object (OBJ), time (TIME), and unknown (UNKNOWN). Each class contains some unique subclasses, so the data set has a total of 84 subclasses. Since there are no instances of UNKNOW type in the question set, we do not consider this class. The question set contains 6260 questions. During the experiment, we took 4960 of them as training samples and 1300 as test samples. The distribution of samples is shown in Table 1.

4.2. Data Processing. After the training sample set and test sample set are determined, these data need to be preprocessed. First of all, the preprocessing of the data includes format conversion, filtering punctuation marks, and special characters. Secondly, for the Chinese question data set, word segmentation is required. In the experiments, we can convert each Chinese question sentence into a sequence of words separated by spaces with the JIEBA word segmentation tool. Then, for Chinese question sentence, stop word removal processing is also needed. Stop words refer to words that often appear in the text without an actual meaning. The stop word list provided by the Harbin Institute of Technology is applied for follow-up experiments. The stop word list includes numeric characters, special characters, and commonly used nonsense words. Finally, the word vector training is carried out on each question sentence.

4.3. Evaluation Standard. In the experiment, we applied the classification accuracy (Acc) to judge the performance of the model, and its formula is defined as follows:

$$Acc = \frac{AccNum}{TNum},$$
 (7)

where AccNum represents the number of questions in the test set that are classified correctly and TNum represents the total number of questions in the test set.

TABLE 1: The number of samples in training set and test set.

Class	DES	HUM	LOC	NUM	OBJ	TIME
Training set	797	328	882	1089	1274	590
Test set	159	177	370	241	198	155

4.4. Experimental Design and Result Analysis

4.4.1. Set Experiment Parameters. We know that traditional CNN training uses a gradient descent method. Generally speaking, although the batch gradient descent method can find the optimal solution, all the total samples are required to participate in the operation when the weight is updated, which will cause a large number of calculations and a slow convergence speed. As for the random gradient descent method, its advantage is that it only needs one sample to participate in the operation each time, and the convergence speed is fast. The disadvantage is that it falls into the local optimal solution easily. So we use the minibatch method for training in the experiments. In this way, the network can speed up the training speed while finding the optimal solution as much as possible to reduce the training loss. We set the minibatch size to 50. When training the word vector, each Chinese question is taken as a line to learn the word vector. And the Skip-gram model of the Word2Vec tool was applied during the process. The dimension parameter d of the word vector was set to 300. In addition, we also set other basic parameters in the experiment, the size of the convolution kernel sliding window is set to 5, and the dropout rate is set to 0.5.

4.4.2. Experimental Results and Analysis.

Experiment 1. Verify that the feature extraction method based on Word2Vec is better than traditional methods in Chinese question classification.

To verify the effect of different feature processing methods on classification, we carried out comparative experiments among the Word2Vec method, the bag of words method, the mutual information, and the IF-IDF method. In the experiment, 2480 pieces of data in the training set were randomly selected as training data, and 650 pieces of data in the test set were used as test data. The classification results are shown in Table 2.

The results in Table 2 prove that, compared with the traditional feature extraction methods, the method based on the Word2Vec has a higher accuracy rate. The main reason is that the feature vectors obtained by training and learning with word vectors can overcome the problem of data feature sparseness in traditional feature training methods. It helps alleviate the dimensionality disaster and high computational complexity. In addition, it can also help improve the accuracy of question classification and reduce the computational complexity of the model.

Experiment 2. Verify that the CNN based on Bagging integrated classification is better than traditional machine learning methods and CNN in Chinese question classification.

Table 2: Classification results based on different feature training methods.

Method	DES	HUM	LOC	NUM	OBJ	TIME	Total
Bag of words	72.87	56.74	64.31	75.16	64.25	60.48	65.67
Mutual information	70.43	54.84	60.66	72.57	62.40	58.16	63.26
IF-IDF	73.95	59.57	64.41	76.35	66.53	62.36	67.82
Word2Vec	79.32	65.95	73.64	84.53	73.06	69.57	74.71

To verify the performance of the W2V + B-CNN model in the application of question classification, comparative experiments were conducted with traditional machine learning methods such as Bayes and SVM, as well as deep network models such as CNN and W2V + CNN. The comparative results are illustrated in Table 3.

From Table 3, the following conclusions can be drawn:

Conclusion 1. The results of Table 3 illustrate that the W2V + B-CNN model proposed in the paper has the best classification effect. Compared with traditional machine methods and CNN models, the W2V + B-CNN model has significantly improved classification accuracy both in the major class and in each subclass. Finally, the experimental results prove the effectiveness and feasibility of the W2V + B-CNN model in question classification.

Conclusion 2. In Table 3, we can also see that the question classification methods based on the CNN model are better than the machine learning method on the task of question classification. It shows that, for the distributed feature representation of word vectors in the task of Chinese question classification, using machine learning methods to construct and match features is not as advantageous as using CNN. The main reason is that the CNN can reduce the parameters of the model through local perception and weight sharing, thereby effectively reducing the complexity of the model. Meanwhile, the classification accuracy of the CNN models is higher than other machine learning models, which shows that CNN performs very well in the classification of Chinese questions.

Conclusion 3. We can also draw the conclusions from Table 3 that, compared with the CNN model, the classification accuracy of W2V + CNN model and W2V + B-CNN in each class has been improved. What is more, the classification accuracy of the W2V + B-CNN model has improved significantly. The reason is that, because the question contains fewer words, the problems of too high feature latitude and sparse feature vector data will occur when the traditional vector space model is used. Moreover, since the correlation between words and the position information of the words in the sentence are not considered, these factors will affect the accuracy of question classification, and the classification effect of CNN is not very satisfactory. Therefore, compared with SVM and Bayes learning methods, the

TABLE 3: Comparative experimental results of different models.

Models	DES	HUM	LOC	NUM	OBJ	TIME	Total
Bayes	85.41	69.28	75.63	87.82	77.58	73.14	78.20
SVM	86.20	68.74	76.48	88.24	77.82	74.15	78.73
CNN	84.65	71.38	78.82	87.64	78.43	74.48	80.81
W2V + CNN	86.46	78.54	79.23	89.17	79.35	77.79	82.62
W2V + B- CNN	88.57	80.24	84.72	92.47	83.18	82.74	85.67

accuracy of the CNN model has a relatively small improvement. Combining the characteristics of short texts, we took the Word2Vec tool to map the features of the words to a certain dimension, then organize the question sentences into the form of a two-dimensional matrix similar to the image, and use it as the input of the CNN model. The advantage is that it solves the data sparsity problem in traditional question classification methods and greatly improves the classification accuracy.

Conclusion 4. Comparing the W2V + CNN model with the W2V + B-CNN model, it can be found that the W2V + B-CNN model can further combine the advantages of the Bagging algorithm and has a greater degree of improvement in classification accuracy than the W2V + CNN model.

In summary, the W2V + B-CNN model proposed in this paper combines the idea of integrated classification with CNN. In this way, the CNN can be used to extract the potential features of the natural question, and the Bagging algorithm can also be utilized to optimize the classifier in the new model. It not only strengthens the positive influence of more complete data on classification but also weakens the negative influence of noisy data, thereby greatly improving the classification performance of the algorithm.

5. Conclusion

Question classification is a key link in the QAS, and its classification performance has an important impact on subsequent document retrieval and answer extraction. It can be said that, as an important submodule of the QAS, the importance of question classification is self-evident. Therefore, to optimize the performance of question classification, we conducted a certain research on the application of CNN and proposed a new model based on the Bagging integrated classification called W2V + B-CNN. Firstly, the model uses Word2Vec for word vector training and then uses the convolutional layer and pooling layer of the CNN for feature extraction and feature selection. Finally, in the integrated classification layer, the Bagging algorithm with better classification performance was used to replace the Softmax classifier for feature classification. On one hand, the new model can apply the CNN to extract the potential features of natural questions. On the other hand, it can also take advantage of integrated learning for feature classification. In the end, to fully verify the feasibility and effectiveness of the W2V + B-CNN model applied to question

classification, we conduct comparative experiments with traditional machine learning algorithms and CNN. The results prove that, compared with other algorithms, the improved W2V+B-CNN model has higher classification accuracy and better question classification performance.

Data Availability

The labeled data sets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Modern Educational Technology Research in Jiangsu Province of China Foundation Project under Grant 2019-R-77164.

References

- [1] B. Ojokoh and E. Adebisi, "A review of question answering systems," *Journal of Web Engineering*, vol. 17, no. 8, pp. 717–758, 2019.
- [2] K. Aurangzeb, Z. A. Muhammad, and A. Hussain, "A rule-based sentiment classification framework for health reviews on mobile social media," *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 6, pp. 1445–1453, 2017.
- [3] P. Chesley, B. Vincent, and L. Xu, "Using verbs and adjectives to automatically classify blog sentiment," *Training*, vol. 580, no. 263, pp. 233–235, 2006.
- [4] E. Boiy and M.-F. Moens, "A machine learning approach to sentiment analysis in multilingual Web texts," *Information Retrieval*, vol. 12, no. 5, pp. 526–558, 2009.
- [5] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6527–6535, 2009.
- [6] L. Wang, H. Zhang, D. Q. Wang, and J. Huang, "Chinese question classification based on semantic gram and SVM," *Computer Science-Technology and Applications*, vol. 111, no. 1, pp. 432–435, 2009.
- [7] J. Silva, L. Coheur, A. C. Mendes, and A. Wichert, "From symbolic to sub-symbolic information in question classification," *Artificial Intelligence Review*, vol. 35, no. 2, pp. 137–154, 2011.
- [8] E. Bisbal, D. Tomás, L. Moreno, J. L. Vicedo, and A. Suárez, "A multilingual SVM-based question classification system," in A Multilingual SVM-Based Question Classification System, pp. 806–815, no. 3789, Springer, Berlin, Heidelberg, 2005.
- [9] G. Wang and S. Y. Shin, "An improved text classification method for sentiment classification," *Journal of information* and communication convergence engineering, vol. 17, no. 1, pp. 41–48, 2019.
- [10] K. L. Jia, X. Z. Fan, and J. Z. Xu, "Classification of Chinese questions based on KNN," *Microelectronics & Computer*, vol. 25, no. 1, pp. 156–158, 2008.
- [11] Y. Rao, H. Xie, J. Li, F. Jin, F. L. Wang, and Q. Li, "Social emotion classification of short text via topic-level maximum

- entropy model," *Information & Management*, vol. 53, no. 8, pp. 978–986, 2016.
- [12] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A convolutional neural network for modelling sentences," *Eprint Arxiv*, vol. 1, pp. 655–665, 2014.
- [13] W. Xin, Y. Liu, and C. Sun, "Predicting polarities of tweets by composing word embeddings with long short-term memory," in Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. Stroudsburg, pp. 1343–1353, ACL, Beijing, China, 2015.
- [14] M. Breja and S. K. Jain, "Analysis of why-type questions for the question answering system," *Communications in Computer and Information Science*, vol. 909, pp. 265–273, 2018.
- [15] X. Li and D. Roth, "Learning question classifiers: the role of semantic information," *Natural Language Engineering*, vol. 12, no. 3, pp. 229–249, 2006.
- [16] P. Schulze, A.-K. Bestgen, R. K. Lech, L. Kuchinke, and B. Suchan, "Preprocessing of emotional visual information in the human piriform cortex," *Scientific Reports*, vol. 7, no. 1, Article ID 9191, 2017.
- [17] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [18] Y. Tian, "Artificial intelligence image recognition method based on convolutional neural network algorithm," *IEEE Access*, vol. 8, no. 99, pp. 125731–125744, 2020.
- [19] A. V. Poliyev and O. N. Korsun, "Speech recognition using convolutional neural networks on small training sets," *IOP Conference Series: Materials Science and Engineering*, vol. 714, no. 1, Article ID 012024, 2020.
- [20] Y. Kim, "Convolutional neural networks for sentence classification," 2014, https://arxiv.org/abs/1408.5882.
- [21] D. Zhang, L. Tian, M. Hong, F. Han, Y. Ren, and Y. Chen, "Combining convolution neural network and bidirectional gated recurrent unit for sentence semantic classification," *IEEE Access*, vol. 6, no. 99, pp. 73750–73759, 2018.
- [22] P. Le and W. Zuidema, "The forest convolutional network: compositional distributional semantics with a neural chart and without binarization," in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1155–1164, Lisbon, Portugal, September 2015.
- [23] Y. Nan, P. Da, and M. Zhang, Stance Detection in Chinese MicroBlogs with Neural Networks, pp. 893–900, Springer International Publishing, Berlin, Germany, 2016.
- [24] S. Yilmaz and S. Toklu, "A deep learning analysis on question classification task using Word2vec representations," *Neural Computing & Applications*, vol. 32, no. 7, pp. 2909–2928, 2020.
- [25] Y. Bengio, R. Ducharme, and P. Vincent, "A neural probabilistic language model," *Journal of Machine Learning Research*, vol. 3, no. 2, pp. 1137–1155, 2003.
- [26] T. Mikolov, K. Chen, and G. Corrado, "Efficient estimation of word representations in vector space," *Computer Science*, vol. 2013, pp. 1–12, 2013.
- [27] Y. Goldberg and O. Levy, "Word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method," pp. 1–5, 2014, https://arxiv.org/abs/1402.3722.