ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

The Teaching Value of Traditional Culture of Chinese Language and Literature under Big Data Analysis and Its Application in Vocational Education

Wei Li[1,*]

[1] School of International Communication, Hunan Mass Media Vocational and Technical College, Changsha, Hunan, 410100, China

SUMMARY: *In this paper, the traditional culture teaching value text data is first processed by Chinese word segmentation and de-duplication, and further converted into word vectors using the DSG algorithm, and substituted into CNN classifiers as inputs to finally complete the design of recognition algorithms for traditional culture teaching value, with a view to increasing the importance of people's attention to the teaching value of traditional culture in Chinese Language and Literature. Based on this, the BiLSTM-Attention combined algorithm was used to conduct relation extraction on the traditional Chinese language and literature cultural word vectors. Thus, a traditional culture domain relation extraction algorithm based on BiLSTM-Attention was established. Subsequently, an intelligent response model for traditional culture was designed through the template matching method. Through this model, traditional culture can be integrated into higher vocational education, thereby achieving the goal of coordinated development of traditional culture inheritance and vocational education. The average accuracy of the traditional culture intelligent answer model is 0.9196, which means that the traditional culture intelligent answer model has excellent performance and can meet the current needs of vocational education in colleges and universities. In addition, the quantitative values of humanistic literacy and professional skills of the students in the experimental group are larger than those of the control group, with values ranging from 3.401 to 3.993, which indicates that the integration of traditional culture into vocational education through the traditional culture intelligent answer model can fully mobilize the students' enthusiasm and interest in learning, and then make the students' level of humanistic literacy and professional skills improve significantly, which is a promising significance for the common development of traditional culture of Chinese language and literature and vocational education in colleges and universities. It has the significance of promoting.*

KEYWORDS: *DSG algorithm; CNN classifier; BiLSTM-Attention; language and literature; vocational education*

1 Introduction

Chinese language and literature, not only is the concentration of language and writing, but also is the key source of Chinese spirit, values and cultural identity. Chinese language and literature plays a major role in the transmission of spirit and values, both as a symbol of spirit and a manifestation of cultural self-confidence [1, 2]. In the course of history, these language and literature texts have passed on the core values of benevolence, righteousness, courtesy and

*18008463868@163.com
https://doi.org/10.65102/is2026634
Published: 30 April, 2026

wisdom, harmony and stability honored by the Chinese people from generation to generation through the transmission of words and ideas, shaping a highly cohesive spiritual outlook. These spiritual values have become the spiritual pillars for the Chinese people to overcome difficulties and pursue progress, and are also important resources for value guidance in modern society [3, 4]. In contemporary times, re-exploring and promoting these traditional cultural meanings can help enhance the sense of national cultural identity, cohesion of social consensus, and enhance the national cultural soft power [5, 6]. In the context of the rapid development of information technology, the forms of dissemination of traditional culture have become increasingly diversified. However, in this process, it also faces the risk of misinterpretation such as “Western-centered”, “fragmentation” and “superficiality” [7, 8]. In addition, the values of modern society have gone through multiple changes, which puts new demands on traditional culture: respect for history, inheritance of excellent culture, and at the same time combining it with contemporary values to meet the social needs of the new era [9, 10].

With the development of globalization and the change of students' ability demand, the status of vocational education has been rising, and the cultivation of students' comprehensive ability has become the focus [11, 12]. And education as an important carrier of inheriting traditional culture, the traditional culture of Chinese language and literature has an important teaching value. On the one hand, students deeply understand the connotation of traditional culture and the values at that time through reading literature and teachers explaining literature, improve their own humanistic literacy, strengthen students' awareness of cultural dissemination and cultural pride, and promote cultural dissemination [13-15]. On the other hand, combined with the background of the information age and the needs of cross-cultural communication, it promotes students to draw inspiration from literary works and recognize cultural differences, which helps to cultivate students' innovative ability and cross-cultural competence [16, 17]. Therefore, under the background of globalization, mining and promoting the teaching value of traditional culture in Chinese language and literature with the help of information technology and applying it to vocational education not only helps to enhance the sense of national cultural identity, but also helps to build a discourse system with Chinese characteristics, so as to show the profound heritage and contemporary value of Chinese culture in the international arena.

In this paper, with the technical support of text mining, the text data of the teaching value of traditional culture of Chinese language and literature are collected, which are processed by Chinese word division and de-duplication, so that the collected data are free of interfering information, which ensures the rigor of the research results. Immediately, the DSG algorithm was used to convert the above text data into recognizable word vectors, and the word vectors were used as inputs to be substituted into the CNN classifier, thus obtaining the recognition algorithm of the teaching value of traditional culture of Chinese language and literature. On this basis, the BiLSTM-Attention combination algorithm is utilized to extract domain relations from the word vectors, which in turn composes the traditional culture domain relation extraction algorithm, and also completes the design of the traditional culture intelligent answer model, through which the vocational education incorporating traditional culture is carried out, which is more conducive to the promotion of the synergistic development of traditional culture and vocational education in the field of big data perspective.

2 Research Methodology

2.1 Algorithm design for recognizing the teaching value of traditional culture

As far as the teaching value of traditional culture of Chinese language and literature is

concerned, through the design of the classification algorithm for the teaching value of traditional culture, it is more conducive to increasing the importance of students' traditional culture of Chinese language and literature by correctly grasping their attitudes towards traditional culture of Chinese language and literature, which in turn will promote the formation of students' correct values. The following is the process of designing the identification algorithm for the teaching value of traditional culture.

2.1.1 Data acquisition

Since this paper is a study of the traditional culture of Chinese language and literature, the experimental data are crawled using text mining technology in big data analysis methods. There is a large amount of dirty data in the raw data returned from crawling, including lost data, inconsistent data, abnormal data, duplicate data, and invalid data. In order to improve the usability of the research data, it is necessary to pre-process the raw data.

2.1.2 Chinese word splitting

In this paper, we use Python Chinese word processing package “jieba” for word processing. jieba has three modes: precise mode, full mode and search engine mode. Among them, the precise mode is in accordance with the word separation algorithm will be a complete sentence precise separation; full mode is to get the sentence can be composed of all the words of the word, this method is fast but can not solve the ambiguity; search engine mode is in the precise mode on the basis of the results obtained, and then cut the long words, improve the recall rate.

The expression of Chinese language and literature traditional culture is basically word-based, so in this paper, the search engine mode is firstly selected to establish the element set one_seg1 by splitting the words of Chinese language and literature traditional culture, and then the Chinese language and literature traditional culture is split word by word to establish the element set, and the two are combined to establish the element set one_seg2.

2.1.3 De-duplication of words

De-deactivation means that if a word in the result after participle appears in the deactivation word list, it can be eliminated. Due to the specificity of the textual data of Chinese language and literature traditional culture, auxiliaries such as “of”, “had”, prepositions such as “when”, “due to”, “because of”, and pronouns such as “I”, ‘you’, and “he” seldom appear, and even if they appear in traditional culture, it is hard to say that they do not contain a large amount of information, and the grammatical structure of traditional culture is not as strict as that of modern Chinese language, therefore, if certain words are removed hastily, it may have a great effect on the categorization effect. Therefore, in this paper, only the punctuation marks and spaces contained in the element sets are removed and updated to one_seg1 and one_seg2 element sets.

2.1.4 Extracting word vectors

Word vectors are distributed low-dimensional real vectors that are generated by training a large-scale corpus that includes a high vocabulary with a neural network language model.The advantage of word vectors is that they take into account synonyms and related words. Traditional text representation methods use vector space models, but the shortcomings of this method are very obvious not only easy to lead to dimensionality disaster, feature sparsity, too much computation, but also does not take into account the existence of synonyms and related words. In this regard the word vector training algorithm used in this paper is the Directional Skip-Gram (DSG) algorithm released by Tencent at the end of 2018.The DSG algorithm adds consideration of the relative positions of word pairs to the commonly used word vector training

algorithm Skip-Gram to improve the accuracy of the semantic representation of word vectors.

The structure of the Skip-gram model is shown in Fig. 1, the Skip-gram model is the same as the CBOW model in that there is no hidden layer. And the difference with CBOW is that the Skip-gram model predicts the context by the target word. The optimization objective for the whole corpus is maximization:

$$\sum_{(w,c)\in D}\sum_{w_j\in c}\log P\left(w\,|\,w_j\right) \tag{1}$$

When the pre-training corpus data is large, Skip-gram works a little better than CBOW, so Skip-gram is more widely used.

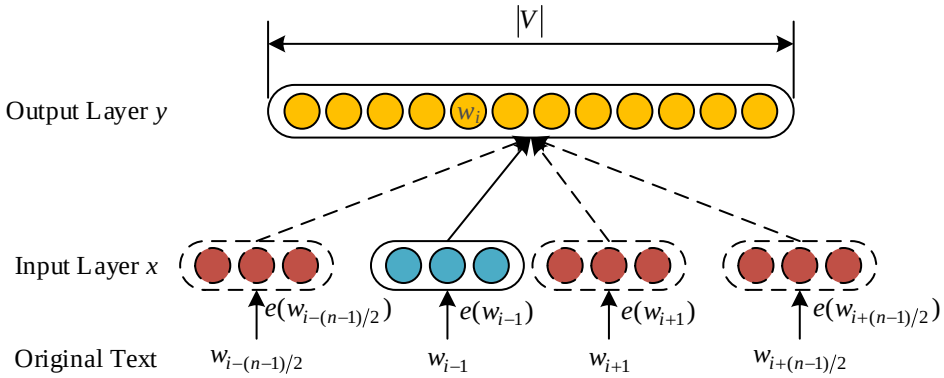


Figure 1: The structure of the Skip-gram model

The model suggests that word order is an important factor influencing word vectors in traditional Chinese language and literature, and that a word is associated with the words to its left or right. Given a word, it is helpful to identify the words to its left or right. In view of this, a softmax function is proposed:

$$g\left(\delta_{w_{t+i}}^{T}v_{w_t}\right)=\frac{\exp\left(\delta_{w_{t+i}}^{T}v_{w_t}\right)}{\sum_{w_{t+i}\in V}\exp\left(\delta_{w_{t+i}}^{T}v_{w_t}\right)} \tag{2}$$

to measure how a context word w_{t+i} is related to the context to its left or right, by introducing a new vector δ for each w_{t+i} to represent its position relative to w_t.

2.1.5 CNN Classifier

The CNN structure used in this paper consists of an input layer, two convolutional layers, two fully connected layers and an output layer, and the CNN structure is shown in Figure 2. The input of the input layer is the 1*1000 -dimensional Chinese language and literature traditional culture word vector extracted by the DSG method. In the first layer of convolution, one-dimensional convolution is used with 32 1*1000 convolution kernels. The output of the first layer convolution is 32 1*1 features. The second layer convolution is a one-dimensional convolution using 32 1*32 convolution kernels. The output of the second layer convolution is 32 1*1 features. The fully-connected layer consists of two fully-connected layers with 64 and 32 neurons respectively, and the activation function is the relu function. A dropout function with a ratio of 0.5 is used. Finally, the output layer undergoes a softmax function to categorize the output into four classes.

The specific formula for the convolutional neural network is as follows:

$$v_j^l = f\left(\sum_{i \in M_j} d_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{3}$$

where v_j^l is the output of the l convolutional layer, M_j denotes the input vector, d_i^{l-1} represents the ith input, f denotes the activation function, and b_j^l denotes the bias term of the lth layer after convolution by the convolution kernel.

In the convolutional layer, d_i^{l-1} is used as the input to be convolved with $i*k$ convolutional kernels, i.e., k_{ij}^l, and finally the output is made by the activation function.

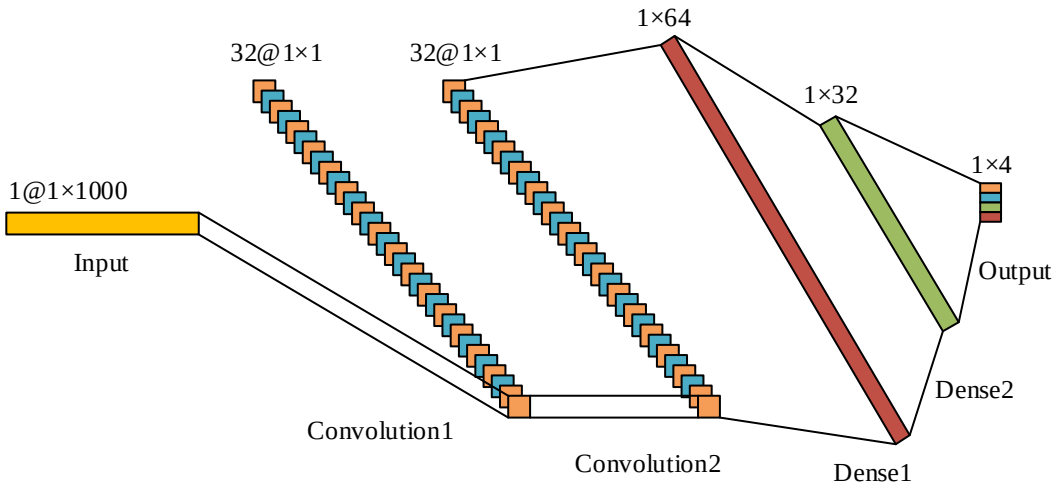


Figure 2: Diagram of CNN structure

2.2 Application of traditional culture in vocational education

Through the design of the identification algorithm of the teaching value of traditional culture, to enhance people's attention to the teaching value of the traditional culture of Chinese language and literature, and to provide a theoretical basis for the integration of the traditional culture of Chinese language and literature into vocational education. On this basis, the BiLSTM-Attention combination algorithm is used to extract the relationship between the word vectors of traditional culture of Chinese language and literature, which establishes a relationship extraction algorithm based on the traditional culture domain of BiLSTM-Attention, and then the intent of the sentence and the question are identified using the template matching method, which finally completes the design of the traditional culture intelligent answer model. Finally, the model is used to integrate traditional culture into vocational education in colleges and universities, so as to realize the goal of synergistic development of traditional culture inheritance and vocational education.

2.2.1 Relational extraction algorithm for traditional cultural domains

In order to compensate for the loss of information in BiLSTM networks due to long distance propagation, an Attention layer is accessed behind BiLSTM. Each step of Attention does not depend on the computation of the previous step and can be computed to the global dependency, thus the combination of Attention and BiLSTM can be applied to the task of relation extraction

to improve the effect of relation extraction in traditional culture domain. Figure 3 shows the schematic structure of the Attention mechanism added after the Bi-LSTM network. The a_{di} represents the probability of the attention weight of the i th node to the d th node.

$$e_{di} = w_1 \tanh\left(w_2 \overrightarrow{h_t} + w_3 \overleftarrow{h_t} + b\right) \tag{4}$$

$$a_{di} = \frac{\exp\left(e_{di}\right)}{\sum_{j=1}^{T} \exp\left(e_{dj}\right)} \tag{5}$$

In the above equation w_1, w_2, and w_3 are the weight matrices, $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ are the forward and backward outputs of the BiLSTM network respectively, and b is the bias term. By splicing through the fully connected layer with the activation function of tanh, and then do the product with the defined weight matrix. Then the similarity coefficients corresponding to the current node and all encoded moments are calculated, while SoftMax normalization is performed. Then the state accumulation of all encoded moments is obtained in the form of weighted summation to get the context vector and finally the output of the current decoded moment. The purpose of this step is to take into account the influence of the current moment in addition to the influence of other encoded states as well.

$$H = \sum_{i=1}^{m} a_{di} * h_i \tag{6}$$

In the above equation H represents the combination of features $H = \{h_1, h_2, h_3, \cdots, h_n\}$ obtained by Attention computation, and n represents the number of sentences.

In this section a SoftMax classifier is used to predict the relations in the sentence S from the set Y of types as labels $\hat{y}$. This classifier takes as input the state h^* of the hidden layer.

$$\hat{p}\left(y \mid S\right) = soft\max\left(W^s h^* + b^s\right) \tag{7}$$

$$\hat{y} = \arg\max \hat{p}\left(y \mid S\right) \tag{8}$$

The loss function of the classifier is the negative log-likelihood equation for the correct type label $\hat{y}$:

$$J\left(\theta\right) = -\frac{1}{m} t_i \log\left(y_i\right) + \lambda \left\|\theta\right\|_F^2 \tag{9}$$

In the above equation t represents the input labeled classes, y represents the probability predicted by the SoftMax classifier corresponding to each class, and m is the sum of the classes. Here L2 regularity is used to prevent model fitting and λ is the hyperparameter of L2 regularity.

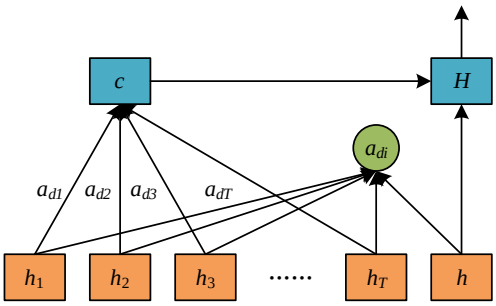


Figure 3: Attention Mechanism model

2.2.2 Intelligent Answer Modeling for Traditional Culture

Intelligent question and answer model, as a classic application of knowledge graph, allows students to talk with machines through natural language to achieve the purpose of human-computer interaction, which is more conducive to attracting the interest of students, and the flow of traditional culture intelligent answer model is shown in Figure 4. According to the demand analysis of the traditional culture intelligent answer model, the input of this model is the speech of the question asked by the students in the form of natural language, and the output is the answer synthesized by speech. The process first converts the students' speech into text through a speech conversion tool, and then extracts the key words in the question with the help of the traditional culture domain relationship extraction algorithm. Then the machine learning intent analysis model is used to identify the intent of the sentence and the question by template matching, so that the query statement searches the Neo4j database for the answer to the question, and determines whether to splice the answer or respond to a prompt according to whether there is an answer in the library.

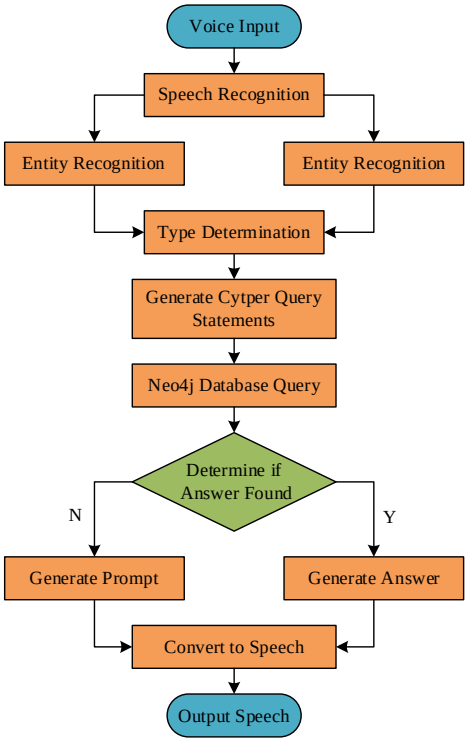


Figure 4: Traditional culture intelligent response model

2.2.3 Application pathways

At present, in the process of integrating traditional culture into vocational education, the problem of insufficient development of digital resources is more prominent, and it is difficult to effectively meet the teaching needs of vocational education. Some vocational colleges and universities have insufficient understanding of the value of traditional culture digital resources, and they still follow the traditional way of presenting cultural content only through flat forms such as text and pictures, which makes it difficult for students to get three-dimensional and visualized learning experience, and the classroom lacks a sense of immersion and interactivity. Through the traditional culture intelligent answer model, the essence of traditional culture is presented in a visualized and interactive form, which significantly improves the effectiveness of traditional culture dissemination. In teaching practice, teachers can rely on the traditional culture intelligent answer model to organically integrate traditional culture into curriculum teaching. This immersive and interactive teaching mode can effectively break through the limitations of traditional one-way indoctrination, fully mobilize the learning enthusiasm of students, make the teaching of traditional cultural knowledge more understandable and deeply rooted in people's hearts, and ultimately realize the synergistic development goal of traditional culture inheritance and professional skills training.

3 Results and analysis

3.1 Exploratory Analysis of Recognition Algorithms

3.1.1 Word Vector Description Statistics

Through text mining technology, we obtained the text data of traditional culture's teaching value. We conducted Chinese word segmentation and removed stop words to obtain the traditional culture's teaching value text data without interference. Then, we used the DSG algorithm to convert the traditional culture's teaching value text data into word vectors. Subsequently, based on structured data, we conducted preliminary descriptive analysis. The descriptive statistics of the word vectors are shown in Table 1. Based on the data in the table, a total of 30 traditional culture's teaching value text word vectors were extracted, which are culture identity, national sentiment, moral cultivation, humanistic quality, spiritual inheritance, value shaping, character cultivation, cultural confidence, historical cognition, aesthetic improvement, national spirit, behavioral norms, intangible cultural heritage inheritance, cultural foundation, moral cultivation, traditional virtues, cultural continuity, thinking inspiration, cultural accumulation, national responsibility, etiquette education, spiritual nourishment, cultural inheritance, personality improvement, classic immersion, cultural consciousness, value guidance, rural sentiment, cultural inheritance, and literacy improvement. The corresponding word frequencies are 85, 85, 85, 85, 85, 85, 85, 79, 72, 72, 68, 66, 64, 60, 58, 58, 58, 56, 53, 53, 47, 46, 46, 44, 44, 42, 38, 37, 35, 30. The top-ranked word vectors were selected as the input for the subsequent classifier, including culture identity, national sentiment, moral cultivation, humanistic quality, spiritual inheritance, value shaping, and character cultivation.

Table 1: Word vector description statistical analysis

word	Frequency	word	Frequency	word	Frequency
Cultural identity	85	National spirit	68	Etiquette education	47
The sentiment of home and country	85	Code of Conduct	66	Spiritual nourishment	46
Moral cultivation	85	Inheritance of intangible cultural heritage	64	Cultural inheritance	46
Humanistic literacy	85	Cultural roots	60	Personality improvement	44
Spiritual inheritance	85	Cultivation of one's character and nature	58	Classic infiltration	44
Value shaping	85	Traditional virtues	58	Cultural awareness	42
Character cultivation	85	The continuation of the cultural context	58	Value guidance	38
cultural confidence	79	Enlightenment of thinking	56	Rural sentiment	37
Historical cognition	72	Cultural heritage	53	Inheritance of civilization	35
Aesthetic improvement	72	National responsibility	53	Improvement of quality	30

3.1.2 Analysis of classifier results

In order to better demonstrate the priority of CNN classifiers, several commonly used classifiers are selected as references, which are logistic regression classifier, plain Bayesian classifier, decision tree classifier, support vector machine classifier, and random forest classifier in order, and the results of different classifiers are represented in the form of classification confusion matrix.

The results of the logistic regression classifier are shown in Table 2, where A, B, C, D, E, F, and G denote cultural identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation, respectively. Through the data performance in Table 2, it can be seen that the classification accuracy of logistic regression classifier for cultural identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation is 0.6, 0.6118, 0.6, 0.5882, 0.5412, 0.5765, 0.6235, and the classification accuracy of logistic regression classifier for the overall classification of traditional culture teaching value is 0.5916.

Table 2: Logistic regression classifier results

Project	A	B	C	D	E	F	G	Accuracy
A	51	2	4	6	8	4	10	0.6000
B	9	52	3	7	9	1	4	0.6118
C	2	7	51	8	5	3	9	0.6000
D	6	2	4	50	9	6	8	0.5882
E	8	7	5	4	46	9	6	0.5412
F	2	4	5	5	9	49	11	0.5765
G	7	5	6	7	1	3	53	0.6235
Mean								0.5916

Using the plain Bayesian classifier to classify the teaching value of traditional culture, the results of the plain Bayesian classifier are shown in Table 3. As can be seen through the data performance in Table 3, the classification accuracy of the Plain Bayesian classifier for cultural

identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation is 0.6471, 0.6588, 0.6353, 0.6706, 0.6235, 0.6588, and 0.6471 in order, while the overall classification accuracy is 0.6487.

Table 3: Naive Bayes classifier results

Project	A	B	C	D	E	F	G	Accuracy
A	55	2	4	6	8	4	6	0.6471
B	9	56	3	5	4	2	6	0.6588
C	3	9	54	2	6	4	7	0.6353
D	10	5	2	57	4	5	2	0.6706
E	7	2	1	6	53	8	8	0.6235
F	1	7	1	6	7	56	7	0.6588
G	2	5	6	4	7	6	55	0.6471
Mean								0.6487

Then the decision tree classifier was used to classify the seven teaching values in traditional culture, and the results of the decision tree classifier are shown in Table 4. According to the size of the data in the table, it can be seen that the decision tree classifier's classification accuracy values for cultural identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation are 0.6941, 0.7176, 0.7412, 0.6824, 0.7294, 0.6941, 0.7059, and the overall classification accuracy of the decision tree classifier for the 7 items of teaching values in traditional culture is 0.7092. The overall classification accuracy is 0.7092.

Table 4: Decision tree classifier result

Project	A	B	C	D	E	F	G	Accuracy
A	59	6	4	2	5	4	5	0.6941
B	4	61	2	3	4	7	4	0.7176
C	2	8	63	1	3	4	4	0.7412
D	2	7	4	58	2	8	4	0.6824
E	3	4	2	3	62	4	7	0.7294
F	4	2	4	3	8	59	5	0.6941
G	2	3	3	4	7	6	60	0.7059
Mean								0.7092

The support vector machine in machine learning algorithm was used to classify the seven teaching values in traditional culture, and the results of the support vector machine classifier are shown in Table 5. The classification accuracy of the support vector machine classifier for cultural identity, national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation is 0.7412, 0.7529, 0.7294, 0.7765, 0.7647, 0.7176, and 0.7412, respectively, which leads to an overall classification accuracy of 0.7462.

Table 5: Support vector machine classifier results

Project	A	B	C	D	E	F	G	Accuracy
A	63	2	5	2	3	5	5	0.7412
B	5	64	4	3	4	3	2	0.7529
C	2	4	62	6	3	4	4	0.7294
D	4	2	4	66	2	3	4	0.7765
E	3	6	4	1	65	4	2	0.7647
F	6	4	2	3	4	61	5	0.7176
G	2	3	3	5	3	6	63	0.7412
Mean								0.7462

Subsequently, the teaching values of Chinese language and literature traditional culture were classified and analyzed by using the random forest classifier, and the results of the random forest classifier are shown in Table 6. Based on the data in the table, it can be seen that the accuracy of the random forest classifier in classifying the seven teaching values of the traditional culture of Chinese language and literature is 0.8000, 0.8235, 0.7882, 0.7765, 0.8118, 0.8353, 0.8118, and the overall accuracy is 0.8067.

Table 6: Random Forest classifier results

Project	A	B	C	D	E	F	G	Accuracy
A	68	1	2	4	3	3	4	0.8000
B	2	70	2	3	1	3	4	0.8235
C	2	2	67	4	3	4	3	0.7882
D	3	2	1	66	6	3	4	0.7765
E	1	5	3	1	69	4	2	0.8118
F	4	1	2	1	2	71	4	0.8353
G	4	1	3	2	3	3	69	0.8118
Mean								0.8067

Finally, the teaching value of traditional culture of Chinese language and literature is classified using the CNN classifier proposed in this paper, and the results of the CNN classifier are shown in Table 7. The size of the data in the table shows that the accuracy of the CNN classifier for the classification of cultural identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation is 0.8941, 0.8824, 0.9294, 0.8941, 0.8941, 0.9412, and 0.8941, respectively, with an overall accuracy of 0.9042, which is higher than that of the logistic regression classifier, plain Bayesian classifier, decision tree classifier, support vector machine classifier, random forest classifier, CNN classifier has priority in the process of classifying the teaching value of Chinese language and literature traditional culture, and the method in this paper can accurately extract the teaching value of Chinese language and literature traditional culture, so that it can better serve the process of vocational education, and then enhance the students' understanding of the traditional culture of Chinese language and literature. The method of this paper can accurately extract the teaching value of Chinese language and literature traditional culture, so that it can better serve the process of vocational education and enhance the students' attention to Chinese language and literature traditional culture, in order to help the students set up the correct values and professionalism.

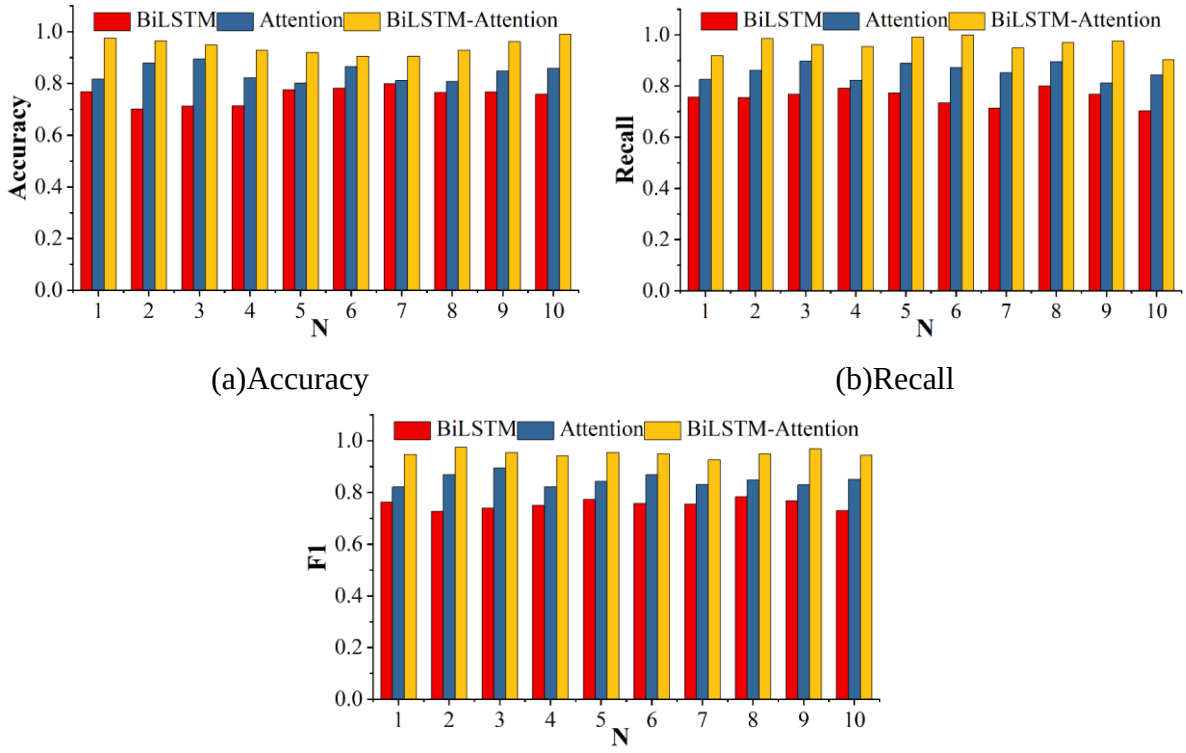
Table 7: CNN classifier results

Project	A	B	C	D	E	F	G	Accuracy
A	76	1	2	1	3	1	1	0.8941
B	2	75	2	1	1	2	2	0.8824
C	0	1	79	1	1	2	1	0.9294
D	1	2	1	76	2	3	0	0.8941
E	1	1	3	1	76	2	1	0.8941
F	1	0	0	2	2	80	0	0.9412
G	1	1	2	2	2	1	76	0.8941
Mean								0.9042

3.2 Analysis of the application of traditional culture in vocational education

3.2.1 Relational extraction algorithm for traditional cultural domains

Using the above proposed BiLSTM-Attention combination algorithm, the traditional cultural domain relations of Chinese language and literature are extracted, and the results of traditional cultural domain relations extraction are shown in Fig. 5, where (a) ~ (c) denote the accuracy rate, recall rate, and F1 value, respectively. From the data size distribution in the figure, it can be seen that when BiLSTM and Attention are used alone to carry out traditional culture domain relationship extraction, the distribution interval of accuracy, recall, and F1 value is 0.712~0.895, and when the BiLSTM-Attention combination algorithm is used to carry out traditional culture domain relationship extraction, the accuracy, recall, and F1 value are improved to 0.902~0.983, indicating that the introduction of Attention on the basis of BiLSTM can improve the effect of relationship extraction in the traditional culture domain of Chinese language and literature, ensure that the subsequent traditional culture intelligent answer model is more in line with the vocational education, and then realize the synergistic development goal of traditional culture inheritance and professional skill cultivation.



(a)Accuracy (b)Recall

(c)F1

Figure 5: The results of relationship extraction in the field of traditional culture

3.2.2 Validation Analysis of Traditional Culture Intelligent Answer Models

In this paper, the traditional culture intelligent answer model is evaluated by the method of manual evaluation, and the traditional culture questions of Chinese language and literature are randomly selected by means of random function, 109 questions are randomly selected, and then the experiment is repeated for 9 times, and the results of the validation and analysis of the traditional culture intelligent answer model are shown in Table 8. According to the results of

the experiments, the average accuracy of the traditional culture intelligent answer model is calculated as 0.9196, from this result, it can be seen that the traditional culture intelligent answer model can accurately answer a part of the questions related to traditional culture knowledge, but there are also a part of the questions that are not answered, which leads to the problem due to two reasons, the first is that the traditional culture knowledge questions in BiLSTM-Attention combination algorithm is extracted incorrectly, so the result from the question query is not consistent with the correct result. The second reason is that there is a deviation in the understanding of the intention of the question, which leads to the wrong direction of the query of the question, e.g., the poems in traditional culture are wrongly understood, which leads to the query of the result of the question not coming out. Overall, the traditional culture intelligent answer model has excellent performance performance, can meet the current needs of vocational education in colleges and universities, and has a facilitating effect on promoting the development of vocational education integrating traditional culture.

Table 8: Model validation analysis results

Experimental round	Experimental results (accuracy)	Experimental round	Experimental results (accuracy)
1	0.9146	6	0.9161
2	0.9216	7	0.9185
3	0.9012	8	0.9276
4	0.9242	9	0.9279
5	0.9248	Mean	0.9196

3.2.3 Application path validation analysis

In this paper, the traditional culture of Chinese language and literature is integrated into the process of vocational education in colleges and universities through the intelligent answer model, in order to prove the effective feasibility of the application path, the Liket five-point scale test is used to verify the effective feasibility of the application path, which tests the students' humanistic literacy and professional skills, and there are 20 items for humanistic literacy and professional skills, respectively, and the scale has an excellent reliability performance, which ensures the credibility of this research. Thirty students were randomly selected from a vocational college and divided into the experimental group and the control group. The experimental group used the intelligent response model to carry out vocational education integrating traditional culture of Chinese language and literature, while the control group used the traditional unidirectional indoctrination method to carry out vocational education integrating traditional culture of Chinese language and literature. The analysis of students' humanistic literacy is shown in Fig. 6, and the analysis of students' professional skills is shown in Fig. 7, in which the black columns indicate the control group and the red the experimental group. Combined with the data performance in the figure, it can be seen that the distribution range of the quantitative values of humanistic literacy and professional skills of the students in the control group is 2.422~2.878, which indicates that the unidirectional instillation method only presents the cultural content in the plane form of text and pictures, and it is difficult for the students to experience the three-dimensional and visualized learning mode, and the classroom lacks interaction and communication, which directly leads to the less obvious effect of the enhancement of the humanistic literacy and professional skills of the students. The distribution range of the quantitative values of humanistic literacy and professional skills of the students in the experimental group is 3.401~3.993, which indicates that teachers can integrate traditional culture into vocational education with the help of the traditional culture intelligent answer model, effectively break through the limitations of one-way classroom knowledge

inculcation, fully mobilize the students' enthusiasm and interest in learning, and make the teaching of traditional culture knowledge more fluent, which further promotes the development of vocational education of integrating Chinese language and literature under the perspective of big data. It further promotes the development of vocational education integrating Chinese language and literature under the perspective of big data, and aims to improve the humanistic quality and professional skills of students in vocational education in colleges and universities.

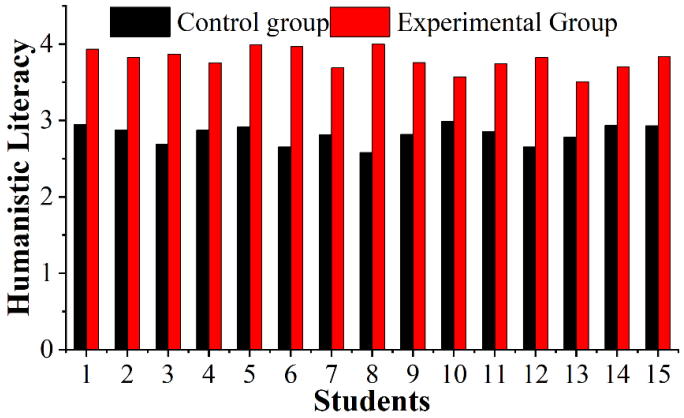


Figure 6: Analysis of Students' Humanistic Literacy

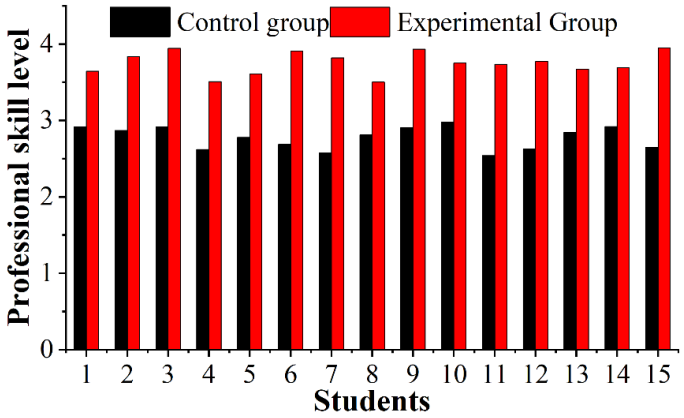


Figure 7: Analysis of Students' professional skills

4 Conclusion

Some colleges and universities do not recognize the value of digital resources of Chinese language and literature traditional culture, and still use the traditional way, which directly leads to the unsatisfactory effect of vocational education. In this regard, the classification algorithm of the teaching value of traditional culture is used to enhance people's attention to the teaching value of traditional culture of Chinese language and literature, and with the help of the intelligent answer model, the traditional culture of Chinese language and literature is applied to vocational education in colleges and universities, with a view to realizing the synergistic development of traditional culture inheritance and vocational education.

(1) The CNN classifier proposed in this paper has a classification accuracy of 0.8941, 0.8824, 0.9294, 0.8941, 0.8941, 0.9412, 0.8941 for cultural identity, family and national sentiment, moral cultivation, humanistic literacy, spiritual inheritance, value shaping, and character cultivation, which is a better prioritization of the CNN classifier compared to the other five, confirming that the The effectiveness of the classification algorithm for the teaching value

of traditional culture is confirmed, and the method can not only enhance people's attention to the teaching value of traditional culture in Chinese language and literature, but also inject fresh blood and vitality into the cause of traditional culture inheritance.

(2) The distribution of the quantitative values of humanistic literacy and professional skills of the students in the control group ranges from 2.422 to 2.878, which indicates that the lack of interaction and communication in the unidirectional knowledge inculcation method directly leads to the less obvious effect of the enhancement of students' humanistic literacy and professional skills. The quantitative values of humanistic literacy and professional skills of students in the experimental group range from 3.401 to 3.993, indicating that teachers can use the traditional culture intelligent answer model to integrate traditional culture into vocational education, fully mobilize students' enthusiasm and interest in learning, and turn vocational education into an activity of mutual communication and learning, which provides theoretical references for the study of vocational education integrating traditional culture of Chinese language and literature under the perspective of big data.

About the Author

Wei Li was born in Changde, Hunan, P.R. China, in 1978. I obtained a master's degree from Fujian Normal University in China. I am currently working at Hunan Mass Media Vocational and Technical College. My main research direction is Chinese language and literature.

References

[1] Sun, L. (2020, October). Application of Big Data Analysis in the Study of the Inheritance of Cultural Confidence in the Teaching of Ancient Literature in Chinese Department. In Journal of Physics: Conference Series (Vol. 1648, No. 2, p. 022156). IOP Publishing.

[2] Zhao, Q., & Castaneda Abdullah, A. Q. (2024). Metaphorical meanings of color symbols in literature. Chinese Semiotic Studies, 20(4), 625-646.

[3] Jiapaer, Y., & Ashan, Z. (2024). Discussion on the Idea Orientation of the Chinese National Community and the Function of National Literature. Journal of Humanities, Arts and Social Science, 8(1).

[4] Fan, Y. (2023). Strengthening the Spirit of Criticism and Enhancing the Sense of Mission: On Constructing the Chinese Form of Literary and Artistic Criticism. Theoretical Studies in Literature and Art, 43(3), 1-14.

[5] Zhao, Z. (2024). Diasporic identity in contemporary sinophone literature: The role of language and cultural elements. Journal of Psycholinguistic Research, 53(1), 13.

[6] Wu, Y. (2025). Mediated Soft Power: A Century of Chinese Science Fiction in Transcultural Trajectory. Critical Arts, 1-18.

[7] Huo, H., Li, Y., & Wang, R. (2019, September). The Predicament of Spreading Chinese Traditional Culture and Methods of Solving the Problem. In 2019 3rd International Conference on Education, Management Science and Economics (ICEMSE 2019) (pp. 480-483). Atlantis Press.

[8] Jinze, L. (2023). Persistence, Optimization and Development: The Dilemma and Path of the Innovation and Development of China Traditional Culture from the Perspective of Media Integration. The Frontiers of Society, Science and Technology, 5(1).

[9] Liu, S. (2023). Analysis of Chinese Excellent Traditional Culture from the Perspective of Cultural Self-confidence. Journal of Humanities, Arts and Social Science, 7(5).

[10] Zou, H., & Liang, Q. (2020). Establishing distinctive theories and methods of design science rooted in Chinese traditional culture. Frontiers of Mechanical Engineering, 15(3), 430-437.

[11] Huang, F. (2025). Exploring a new model of undergraduate vocational education. Vocation, Technology & Education, 2(1).

[12] Daryono, R. W., Ramadhan, M. A., Kholifah, N., Isnantyo, F. D., & Nurtanto, M. (2023). An empirical study to evaluate the student competency of vocational education. International Journal of Evaluation and Research in Education (IJERE), 12(2), 1079-1086.

[13] Sun, S. (2020). Improving Students' Humanistic Quality in the Teaching of Chinese Language and Literature in Vocational Higher Education Institution. Journal of Contemporary Educational Research, 4(7).

[14] Chen, L., & Yuan, M. (2022). Analysis on the cultivation of college students' cognitive psychology of chinese excellent traditional culture by chinese ancient literature. Psychiatria Danubina, 34(4), 1170-1175.

[15] ZHANG, Y. (2023). The Function of Inheriting Traditional Culture in Chinese Language and Literature. DEStech Transactions on Social Science, Education and Human Science, (isss).

[16] Zhang, Y. (2022, July). Cross-cultural literary comprehension: Theoretical basis and empirical research. In Interkulturelles Forum der deutsch-chinesischen Kommunikation (Vol. 2, No. 1, pp. 58-73). De Gruyter.

[17] Geng, B., Huang, T., Jiang, X., Lin, N., Gao, G., & Fan, L. (2021). The analysis of the innovation consciousness of college student entrepreneurs Under the teaching concept of Chinese excellent traditional culture. Frontiers in psychology, 12, 717336.