

Article

Sentiment Analysis of Online Course Evaluation Based on a New Ensemble Deep Learning Mode: Evidence from Chinese

Xiaomin Pu ¹, Guangxi Yan ², Chengqing Yu ², Xiwei Mi ^{3,*} and Chengming Yu ²

¹ College of Information Engineering, Hunan Industry Polytechnic, Changsha 410036, China; pxm0911123@163.com

² School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China; yanguangxi@csu.edu.cn (G.Y.); yuchengqing@csu.edu.cn (C.Y.); 1109150133@csu.edu.cn (C.Y.)

³ School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

* Correspondence: mixiwei@bjtu.edu.cn; Tel.: +86-152-0086-6698

Abstract: In recent years, online course learning has gradually become the mainstream of learning. As the key data reflecting the quality of online courses, users' comments are very important for improving the quality of online courses. The sentiment information contained in comments is the guide of course improvement. A new ensemble model is proposed for sentiment analysis. The model takes full advantage of Word2Vec and Glove in word vector representation, and utilizes the bidirectional long and short time network and convolutional neural network to achieve deep feature extraction. Moreover, the multi-objective gray wolf optimization (MOGWO) ensemble method is adopted to integrate the models mentioned above. The experimental results show that the sentiment recognition accuracy of the proposed model is higher than that of the other seven comparison models, with an F1score over 91%, and the recognition results of different emotion levels indicate the stability of the proposed ensemble model.



Citation: Pu, X.; Yan, G.; Yu, C.; Mi, X.; Yu, C. Sentiment Analysis of Online Course Evaluation Based on a New Ensemble Deep Learning Mode: Evidence from Chinese. *Appl. Sci.* **2021**, *11*, 11313. <https://doi.org/10.3390/app112311313>

Academic Editor: Valentino Santucci

Received: 24 October 2021

Accepted: 25 November 2021

Published: 29 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: two-channel word vector; deep features mining; attention; multi-objective optimization ensemble

1. Introduction

With the integration of internet technology into education and the expansion of the demand for online teaching in schools during the epidemic [1], online learning has been widely promoted to school education, professional education, party member education, and other educational learning classes. Thus, diversified online courses, such as MOOC, SPOC, and NetEase Cloud Classes, have emerged. These online courses focus on providing personalized learning programs for different learners, and their convenient communication learning methods and advanced education concepts have triggered educational and teaching reforms at all levels of learning [2]. Thanks to the development of new technologies, such as artificial intelligence, online interaction, and remote control, the online teaching experience has been greatly enriched. However, users also put forward higher requirements for online courses in terms of rich content, depth of theory, and quality of service. Some long-standing problems in online education, such as uneven quality of online courses and difficulty in guaranteeing learning effectiveness, also need to be solved [3]. The large scale of users has produced a large amount of online education data [4]. How to extract effective information for curriculum improvement and provide personalized learning suggestions for different users has attracted extensive attention from scholars and online education enterprises at home and abroad.

Users in online learning commonly expound their views in the comments and evaluation. Through comments, users can realize the exchange of knowledge to understand or course evaluation. Potential users can learn about the course content and other user experience. Besides, the platform can also improve and enrich course teaching content by

students' and teachers' comments. It can be found that online education is different from other commodities or products. Its user comments include understanding of knowledge, opinions on arguments, and evaluation of courses, among which the evaluation of course content and online teaching methods is of great significance to the improvement of online education [5]. As an important basis to reflect students' acceptance of courses and online teaching quality, the text data of online course evaluation has a strong emotional color. The analysis of students' evaluation of online courses can provide guidance for further optimization of online education quality. The text sentiment analysis is conducive to mining effective suggestions on online education and is an important way to improve online course satisfaction and online education quality.

Text sentiment analysis is an important branch of comment text mining. Its main task is to recognize the emotional tendency and emotional degree of subjective text, and then complete the induction and classification of the text content, so it is also called opinion mining [6]. At present, the mainstream sentiment analysis methods can be divided into the sentiment dictionary method [7], traditional machine learning method, deep learning, and other methods. As the most classical text sentiment analysis method, the sentiment dictionary method has the characteristics of simple operation, but the method is limited in the establishment of sentiment dictionary, thesaurus scope, accuracy, and other aspects [8]. Traditional machine learning methods, such as the K-nearest neighbor algorithm (KNN), decision tree (DT), and support vector machine (SVM), can ensure the accuracy of sentiment classification to a certain extent [9], but most of the above methods rely on manually labeled training sets and features extracted by established methods. The accuracy of training set labeling and the effectiveness of features largely determine the accuracy of traditional machine learning methods. In recent years, the application of deep learning methods in text emotion analysis has become a research hotspot, such as convolutional neural network (CNN), recurrent neural network (RNN), and other deep networks [10]. They have remarkable effects in adaptive text feature extraction. Moreover, combined with word vector representation, attention methods, or other methods, they can achieve high precision text emotion analysis. However, a single model always has its advantages and limitation in terms of accuracy and robustness. On the one hand, each kind of deep network has its unique character and insufficiency. For example, CNN is able to learn the local space characteristics of the text but cannot analyze its temporal characteristics; RNN can realize timing characteristics extraction but cannot obtain long-distance information [11]. On the other hand, different word embedding methods also have their own advantages and weaknesses in the representation of Chinese words. The Word2Vec method relies on the semantic features of local corpus to establish word vectors, while the GloVe method pays more attention to the co-occurrence of statistical features of words in the corpus [12]. How to combine the advantages of different text sentiment analysis methods to further improve the stability and accuracy of the model is the current research focus of text sentiment analysis, and also the focus of this paper.

The structure of this paper is organized as follows: Section 2 mainly introduces the related work of emotional text analysis and the contribution of the paper; Section 3 introduces in detail the model method and ensemble framework proposed in this paper. In Section 4, comparative experiments are carried out to verify the performance improvement of the proposed model in terms of accuracy and validity. Finally, Section 5 summarizes and prospects the main contributions of this paper and the improvement direction of sentiment analysis.

2. Related Works and Contribution

2.1. Word Vector Generation

Word vector representation is the process of digitization of emotional text content. By using high-dimensional vectors to replace different words, the text can be efficiently processed by machine learning, deep learning, and other model algorithms. The word vector representation method has been widely concerned as the first step of in-depth text analysis,

and many word embedding methods, such as shallow text representation method [13], GloVe [14], Word2Vec [15], and FastText [16], appear. Shallow text representation is a relatively traditional word embedding method, which mainly relies on one-hot encoding to transform the text into high-dimensional text representation vectors [17]. However, due to the sparsity of the generated vector, this method cannot meet the input form of most machine learning algorithms. In order to solve the above problems, most word embedding methods mainly rely on neural networks to learn and extract the deep semantic information of shallow text representation and then provide the low-dimensional word vector containing the deep text information.

The Word2Vec word embedding method is a mainstream word vector representation method based on this idea, which takes the high-dimensional sparse one-hot encoding vector as input and transforms it into low-dimensional vectors by using a neural network. This method uses the vector distance to measure semantic similarity between texts in the training process and has attracted great attention and development since it was proposed in 2013. Mikolov et al. proposed the word bag model (CBOW) and skip-gram model in the training process [18]. The former determines the semantic meaning of a word according to the continuous text of the word, while the latter predicts the word vector of the word by using adjacent words. The above two methods make Word2Vec suitable for large-scale corpus training. Naderalvojud et al. proposed an improved word embedding method, which added emotional information encoding to the word vector pre-trained by Word2Vec, making the word vector suitable for emotional text analysis [19]. Onan combined word embedding methods with cluster analysis, and it was proved that the text analysis accuracy was improved by the method [20]. Muhammad et al. utilized Word2Vec and LSTM to analyze hotel reviews and found that the recognition precision rate could reach 85% [21]. The Word2Vec method has been studied by domestic and foreign scholars, and word vectors extracted in most scenarios have good performance. However, Word2Vec fixed the size of the text input window in the training process and inevitably ignored the word co-occurrence statistics of the whole text despite in-depth analysis of local expected features.

Jeffrey et al. [22] proposed the GloVe model in order to make better use of the global word co-occurrence statistical information of the corpus. In this model, the co-occurrence matrix is constructed to count the number of times that each word appears in the window and adjacent words appear in the window. The model can extract the global context information features of words using the information of adjacent words. The GloVe word vector representation method is widely used in the field of text analysis due to its global information analysis ability, so it is not selected to attract scholars for in-depth study. Cao et al. proposed a calibration method for the GloVe model to further improve its performance [23]. In this method, the estimated value of the power-law index approaches 1 by power transformation of the co-occurrence logarithm, and the co-occurrence logarithm data after transformation is used to replace the co-occurrence logarithm of GloVe so that the accurate distribution representation of words can be learned. Mehran et al. proposed a new content tree word embedding method into the GloVe model and found that the text classification accuracy of the improved word vector was significantly improved [24].

As the first step of text analysis, word vector representation has attracted great attention. New methods, such as cross-context word embedding (ACWE) [25], adaptive word embedding (L2AWE) [26], and FastText [27], have been proposed. However, single-channel word vector embedding has an inevitable limitation, as it cannot analyze both the global and local information of the text. Therefore, the ensemble learning method is adopted to combine the two types of word vectors to improve the overall accuracy and stability of the model.

2.2. Text Sentiment Analysis Methods

As a hot topic of text analysis, text sentiment analysis has been widely concerned and studied by scholars at home and abroad. While text features are gradually developing

from TF-IDF and other text statistical features to word vector representation, machine learning and deep learning methods are also applied to this field [28]. Compared with traditional sentiment dictionary methods, machine learning and deep learning methods are characterized by simple data processing, high efficiency, and strong universality. As many large corpora can be used for pre-training, there is no need to worry about the inaccuracy caused by the thesaurus scope [29]. The idea of applying machine learning methods to text sentiment analysis is to use text data with emotion labels to train machine learning models, which can realize sentiment analysis of text in test sets after potential rules are obtained. Chintalapudi et al. used the dictionary method and Bayesian algorithm to achieve the emotional classification of medical documents and found that the Bayesian network can reach 80% classification accuracy [30]. Song et al. proposed a new semi-supervised learning framework based on SVM and found that this method has an excellent performance in text sentiment analysis, and has significantly improved performance compared with general dictionary methods and machine learning methods [31]. In order to recognize semantics emotion terms and emojis, Jia proposed a new sentiment classification framework [32]. The method combined the dictionary methods with the word embedding method and reached satisfactory results. Many other machine learning methods, such as KNN and DT, have also been applied in the field of text sentiment analysis. However, since most emotion text data do not have emotion labels, such supervised learning algorithms are greatly limited in the preparation of training sets [33].

In recent years, deep learning methods have been continuously applied in various fields. Deep learning methods focus on deep feature mining and can realize text emotion analysis combined with attention mechanism [34], which has great advantages in the text emotion analysis. Nassif et al. analyzed the performance of CNN and RNN when applied to Arabic subjective sentiment analysis and verified their capacity in natural language processing (NLP) [35]. Liao et al. used a multi-layer graph convolutional network to recognize text emotions, and set different window sizes at each layer to ensure that the network could learn local and global features, thus improving the emotion recognition effectiveness of the graph convolutional network [36]. Mohammad et al. verified the effectiveness of CNN and LSTM in text sentiment analysis and found that the recognition accuracy of the deep learning model was higher than that of the general machine learning model [37]. However, different deep networks have their own characteristics. In the face of complex text emotion information, specific deep networks can only dig local deep features. For example, CNN focuses on local semantic feature extraction, but its pool operation will disrupt the text order, which is not conducive to temporal feature analysis [38]. LSTM can learn certain temporal features, but it is not in-depth enough to analyze local text features [39]. Therefore, this paper adopts ensemble learning to integrate different deep network emotion recognition models to improve the overall accuracy and stability of the model.

2.3. Ensemble Methods

Ensemble learning is an important part of machine learning, which enables the overall model to have higher recognition accuracy and stability by combining multiple models with similar performance [40]. In recent years, many integrated learning frameworks have been proposed, including the bagging method, stacking method, optimization method, etc. Bagging and stacking are suitable for scenarios with more submodels due to their ensemble mechanisms. The bagging method and voting method are commonly used to construct an online comment sentiment classification ensemble model [41]. However, this method is only applicable to the situation where there are many base classifiers; otherwise, the voting method will lead to overall instability of the model. The optimal ensemble method is generally used to combine different models by optimizing the weight of each sub-model, and is more suitable for the case of a small number of sub-models with good performance. In this paper, multi-objective gray Wolf optimization is used to optimize the weight of the base model, and the ensemble of the two-channel word vector representation

method with the CNN and LSTM deep network is realized. The experiments verify that the ensemble model has higher accuracy and stability compared with other methods in emotion recognition of online course comment text.

2.4. Contribution of This Work

This paper proposes an optimized and ensemble hybrid deep network model for sentiment analysis of online learning evaluation texts.

- (1) The two-channel word embedding method is used to represent online course evaluation texts, which can reduce data sparsity and ensure data integrity.
- (2) Feature training was carried out by CNN and BiLSTM, respectively, and emotion analysis of each submodel was completed by combining the attention mechanism. The deep network was able to extract different aspects of text depth features, and the stable emotion recognition effect of the deep network laid a foundation for the optimal ensemble method.
- (3) The weight of each sub-model was trained based on the multi-objective gray Wolf optimization algorithm to obtain the final sentiment analysis results. The experimental results show that this model has high accuracy and high stability for different data sets.
- (4) The model proposed in this paper provides a meaningful reference for the sentiment analysis of online course evaluation. The ensemble deep learning model proposed in this paper is a new framework for text emotion recognition. In addition, compared with the models proposed by 10 other researchers, the proposed ensemble deep learning model adopted in this paper can achieve the optimal recognition accuracy.

3. Methodology

3.1. Model Framework

The premise of extracting useful teaching suggestions is to analyze the emotional tendency and degree of students' comments through text analysis. This paper proposes a novel text sentiment analysis method to analyze the sentiment tendency of curriculum evaluation texts. The proposed hybrid deep neural network model, which is ensemble by Multi-Objective Gray Wolf Optimization, is shown in Figure 1.

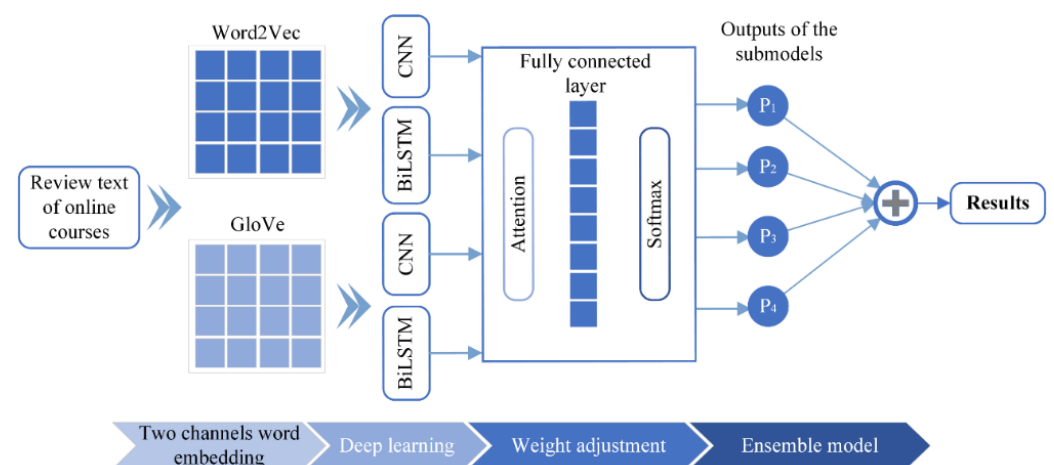


Figure 1. The framework of the proposed model.

Firstly, the two-channel word embedding method is conducted. The Wikipedia text data is utilized to train Word2Vec and GloVe, respectively. The huge amount of data in Wikipedia will help improve the performance of word embedding models. Then, by applying the two well-trained models to the online course comment text, the comment text vector of two types is obtained. Secondly, CNN and bidirectional LSTM are used to analyze the text in the process of deep feature mining. CNN can deeply dig out the local features of the text. However, bidirectional LSTM can deeply analyze the long-distance

temporal sequence information of features, and the obtained text features contain their contextual semantic information. By inputting the text vectors represented by two different word vectors into the above two networks, four different base-depth network models can be obtained. In the emotion analysis part, the local text state vectors obtained by the four base models were input into the attention mechanism, respectively, and the text emotion recognition probability of each base model was obtained after selection by the full connection layer and Softmax. Finally, in the weight optimization part, the weight of each sub-input is trained by gray Wolf optimization, and the weighted value of the recognition probability of each sub-output is taken as the final emotion recognition result of the online education course review text.

3.2. Two Channels of Word Vector Embedding

The neural network model only supports numerical processing, so the digitization of text is the first step of text analysis. In this paper, the two-channel word vector is used to represent the text, as shown in Figure 2. Firstly, the Chinese Wiki corpus is left to jieba word segmentation, which is used for pre-training of the Word2Vec and GloVe word embedding models. Finally, the two word embedding models obtained by pre-training are used to represent the evaluation text, respectively, and two different text representation vector sets are obtained. The sentiment analysis of online course evaluation text using two-channel word embedding can retain the original text information to the maximum extent, which is helpful for the subsequent in-depth text feature mining. During Word2Vec training, the skip-gram algorithm was used to output a 240-dimensional word vector. During Glove training, the number of iterations is set to 50 and a 200-dimensional word vector is output.

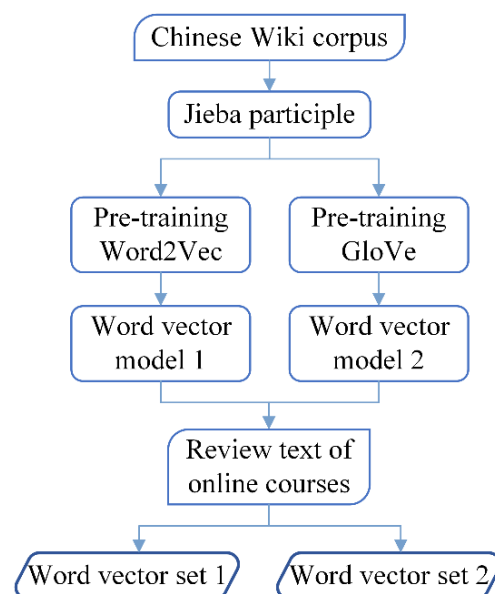


Figure 2. Flowchart of word embedding.

3.3. Deep Networks

CNN and LSTM both have the ability to deeply mine the potential features of text vectors. CNN focuses on the evaluation of text local feature mining and can obtain features between multiple consecutive words. LSTM can capture the temporal and global features of text without gradient explosion or gradient disappearance. Taking the above two models as sub-models, the problem of information loss caused by using two networks successively to process text vector can be avoided. At the same time, more comprehensive text emotion characteristics can be obtained for online course evaluation.

3.3.1. Convolutional Neural Network

The convolutional neural network (CNN) is the most widely used deep neural network. It has a strong feature learning ability and feature representation ability, and has a relatively stable effect in dealing with text classification and emotion recognition [42]. The two text vector data obtained in Section 3.2 were input into the CNN network, respectively, and multiple convolution kernels with different window lengths were used for convolution. Then, the feature matrix obtained by convolution was pooled and the final text feature vector F_C was obtained through repeated iteration [43].

In the process of convolution, n convolution check input text vectors are used for convolution in this paper to learn local text features. Each convolution kernel can obtain the local text feature matrix C [44]:

$$C = \begin{bmatrix} c_{1,1}, c_{1,2}, \dots, c_{1,h} \\ c_{2,1}, c_{2,2}, \dots, c_{2,h} \\ \vdots \\ c_{k,1}, c_{k,2}, \dots, c_{k,h} \end{bmatrix} \in \mathbb{R}^{k \times h} \quad (1)$$

where h is the length of the word vector, and the matrix element C_{ij} can be obtained by the following formula [45]:

$$c_{i,j} = z(\delta + K \times v_{i,i+m-1}) \quad (2)$$

where z is the activation function; $v_{i,i+m-1}$ represents the m input word vectors, which are continuous; δ is the bias value; $K \in \mathbb{R}^{m \times h}$ is the convolution kernel; and RELU is the activation function used in this paper.

3.3.2. Bidirectional LSTM

Bidirectional LSTM (BiLSTM) is improved to make up for the problem that LSTM can only learn the previous information of the text words forward [21]. BiLSTM can completely learn the context information of the text and improve its learning ability of the sequential features of the online course evaluation text. The two text vectors $v_{i,i+m-1}$ obtained in Section 3.2 are input, respectively. The text vector is the splicing of m continuous word vectors. The hidden layer of bidirectional LSTM sets forward and backward propagation layers to learn the context information of the input vector [46]. Each communication layer is connected to the output layer, so that W_f , W_i , and W_o respectively represent the weights of each communication layer to the forgetting door, input door, and output door, then the states of each gate at time t can be calculated according to the following formula [47]:

$$\begin{cases} f_t = F_{\text{sigmoid}}(\lambda_f + W_f \times [y_{t-1}, v_t]) \\ i_t = F_{\text{sigmoid}}(\lambda_i + W_i \times [y_{t-1}, v_t]) \\ o_t = F_{\text{sigmoid}}(\lambda_o + W_o \times [y_{t-1}, v_t]) \end{cases} \quad (3)$$

where F_{sigmoid} represents the activation function of the gates, which is the sigmoid function, y_{t-1} is the outputs at $t-1$. The output of LSTM is calculated as follows [48]:

$$l_t = f_t \odot l_{t-1} + i_t \odot \tanh(\lambda_v + W_v \times [y_{t-1}, v_t]) \quad (4)$$

$$y_t = o_t \odot \tanh(l_t) \quad (5)$$

where l_t represents the status of the LSTM units. Since bidirectional LSTM has two propagation layers and a hidden output, vector h can be obtained at each moment, and the hidden output of bidirectional LSTM can be expressed as follows [49]:

$$F_L = \{h_{f1}, h_{f2}, \dots, h_{ft}, h_{b1}, h_{b2}, \dots, h_{bt}\} \quad (6)$$

where h_f represents the output of forward propagation of the hidden layer, and h_b represents the output of backward propagation of the hidden layer.

3.3.3. Attention Layer

The idea of the attention mechanism is to extract the key information that is consistent with the current purpose through weight adjustment [50]. Weight adjustment by the attention mechanism is beneficial to discover the text feature representation that is important for emotion recognition and then improve the accuracy of emotion recognition. The feature outputs obtained by the two deep neural networks are weighted with the attention mechanism, respectively. Within the length of the statement window m , the new text feature vector u can be calculated as follows [51]:

$$u = \sum_{i=1}^b \beta_i x_i \quad (7)$$

where β_i is the weight of feature vectors, i represents the number of the variables and x_i represents the convolutional features obtained from hidden states or different pooling layers. β_i can be calculated as follows [52]:

$$\beta_i = \frac{\exp(f(\lambda_x + W_x \times x_i))}{\sum_i \exp(f(\lambda_x + W_x \times x_i))} \quad (8)$$

where f is the nonlinear transformation function. W_x and λ_x are parameters. RELU was used as the main nonlinear activation function in this study. Finally, the new text feature vector u was spliced, and the text feature expression u is input into the full connection layer. The softmax function was used for emotion analysis, and the emotion recognition probability value of the text was obtained.

3.3.4. Multi-Objective Gray Wolf Optimization (WOGWO)

In the work, the Multi-Objective Gray Wolf Optimization [53] method is used to carry out weight training on the output probability of the four base models. The F1score and cross-entropy loss are taken as objective functions. The calculation method of F1score will be introduced in Section 4.2, and the cross-entropy loss formula is as follows [54]:

$$L = - \frac{\sum [s \ln \hat{s} + (1 - s) \ln(1 - \hat{s})]}{n} \quad (9)$$

where s represents the true value, \hat{s} is the predicted value, L is the cross-entropy loss, and n is the number of samples.

The specific steps of multi-objective gray Wolf optimization are as follows:

- Step 1: Initialize parameters, including the number of gray wolves, maximum iteration, search scope, and external archive parameters.
- Step 2: Initialize the gray wolves. Firstly, gray wolves were randomly generated, and the constraint conditions were checked. Repeat the above steps until sufficient qualified personnel is produced. Finally, the objective function value is calculated, the non-dominant individual is determined, and the file is updated.
- Step 3: α , β , δ wolves are selected from the archives according to the roulette method, and the remaining wolves are updated according to the positions of α , β , δ wolves to check whether the newly generated wolves meet the constraints. This process is repeated until a sufficient number of qualified gray wolves are produced.
- Step 4: Calculate the objective function, determine the non-dominant gray Wolf, and update.
- Step 5: Repeat Steps 3 and 4 until the end condition is met.
- Step 6: Output the location of the gray wolves in the external archive, the Pareto solution set.

Through the MOGWO algorithm, the ensemble of four base models can be realized, and the advantages of two-channel word vector representation and different depth neural

networks can be used to achieve accurate and stable sentiment recognition of online course evaluation text.

4. Results and Discussion

4.1. Data Description and Preprocessing

During the epidemic period of the past two years, the university's online course platform has been fully improved, and students' requirements on the quality of online courses and online interaction between teachers and students have also been put forward. The source of the text sentiment database adopted in this paper is the relevant evaluation and corresponding scoring of the students at our school on the online teaching content of teachers. The data set of student evaluation texts and scores are exported through the school's online course platform, which can truly reflect the school's online course teaching effect. Each piece of data consists of the following two parts: (1) students' specific evaluation of the quality of the online course and (2) the student's score for the online course is divided into five grades.

Before applying the data to the training model, the reliability and accuracy of the data should be guaranteed. Therefore, first of all, it is necessary to screen and preprocess data, including simplified and traditional style conversion, deletion of too long and too short data, deletion of repetitive text content, deletion of symbols without special functions, etc. In addition, in order to ensure the balance of the sample, it is necessary to equalize the collected data. Table 1 shows the specific construction of the emotional intonation library after final pretreatment and screening:

Table 1. Information of the dataset.

Class Number	Mood	Number of Samples
1	Great satisfaction	4020
2	Satisfaction	4320
3	Neutral	4670
4	Disappointment	3830
5	Great disappointment	3400

After screening and preprocessing, a total of 20,240 tones containing emotion labeling were obtained. Due to sample equilibrium processing, the data volume of each emotion category is about 4000, among which the data volume of emotion is the largest, and the data volume of emotion is very disappointing in the least. In order to fully train the model and establish the classification model with excellent performance, it is necessary to divide the data set into the training set, verification set, and test set. This paper selects 100 samples from each category of data as test sets to evaluate the performance of the model. In addition, 200 samples from each category of data were selected as validation sets to train the parameters related to integrated learning and prevent the model from falling into over-fitting. Finally, the remaining data were used as a training set to train the four base models proposed in Section 2.

4.2. Evaluation Metrics

For the traditional binary classification problem, the samples can be divided into four cases: true cases (TP), false-positive cases (FP), true-negative cases (TN), and false-negative cases (FN). The confusion matrix of the classification results is shown in Table 2.

Table 2. Confusion matrix of classification indexes.

True Label	Predicted Label	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

The text emotion recognition task in this paper is a multi-classification task. In this task, Accuracy, Precision, Recall, and F1 values are commonly used evaluation indexes. Accuracy is the percentage of the total sample that predicts the correct outcome. Accuracy is taken as the most original evaluation index, and its calculation formula is as follows [55]:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (10)$$

Accuracy effectively reflects the probability that all predicted sample values are true. It can be calculated by the following formula [56]:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (11)$$

Recall is the probability that a sample that is actually positive is predicted to be positive. The calculation formula of Recall rate is given as follows [57]:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (12)$$

The F1 value represents the weighted average of Accuracy and Recall. When the numerical difference between Accuracy and Recall rate is large, the F1 value can effectively combine the two indicators. The formula for calculation of the F1 value is as follows [58]:

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

4.3. Comparison Experiments and Results Analysis

4.3.1. Comparison of Different Word Embedding Models

In order to prove that the proposed word embedding model has excellent ability regarding text sequence modeling, the proposed two-channel word embedding model is compared with the traditional word bag model. In addition, in order to prove the advance and advancements of the two-channel model proposed in this paper, the proposed model is compared with the classical Word2Vec and GloVe models. Table 3 shows the recognition results of different models. Figure 3 shows the confusion matrixes of different models. Based on Table 3 and Figure 3, the following conclusions can be drawn:

- (1) Compared with the word embedding modeling algorithm, the traditional word bag model achieves the worst classification result. This fully shows that the word embedding model can effectively extract the depth feature information from the original text data, which effectively optimizes the input of the subsequent classifier of the model. Therefore, the word embedding model has excellent application value in the field of emotion recognition.
- (2) Compared with the classical Word2Vec and GloVe models, the two-channel word embedding model adopted in this paper has a more excellent classification effect. This fully proves that the proposed model has excellent application value in the field of emotion recognition modeling. Two-channel word embedding modeling extracts word vectors from two different angles and transmits the extracted feature information to the classifier. More feature data lay a foundation for further classification.

Table 3. Emotion recognition results of different word embedding models.

Model	Evaluation Metrics	Mood States					Average
		Great Satisfaction	Satisfaction	Neutral	Disappointment	Great Disappointment	
The proposed model	Accuracy	0.9760	0.9600	0.9700	0.9580	0.9720	0.9672
	Precision	0.9783	0.8846	0.9126	0.8692	0.9574	0.9204
	Recall	0.9000	0.9200	0.9400	0.9300	0.9000	0.9180
	F1	0.9375	0.9020	0.9261	0.8986	0.9278	0.9184
Word2Vec	Accuracy	0.9740	0.9320	0.9480	0.9380	0.9600	0.9504
	Precision	0.9780	0.8000	0.8700	0.8286	0.9255	0.8804
	Recall	0.8900	0.8800	0.8700	0.8700	0.8700	0.8760
	F1	0.9319	0.8381	0.8700	0.8488	0.8969	0.8771
GloVe	Accuracy	0.9620	0.9520	0.9420	0.9580	0.9580	0.9544
	Precision	0.9175	0.8585	0.8381	0.8990	0.9247	0.8876
	Recall	0.8900	0.9100	0.8800	0.8900	0.8600	0.8860
	F1	0.9036	0.8835	0.8585	0.8945	0.8912	0.8862
Word bag model	Accuracy	0.9320	0.9260	0.9200	0.9260	0.9440	0.9296
	Precision	0.8511	0.8119	0.7830	0.8058	0.8750	0.8254
	Recall	0.8000	0.8200	0.8300	0.8300	0.8400	0.8240
	F1	0.8247	0.8159	0.8058	0.8177	0.8571	0.8243

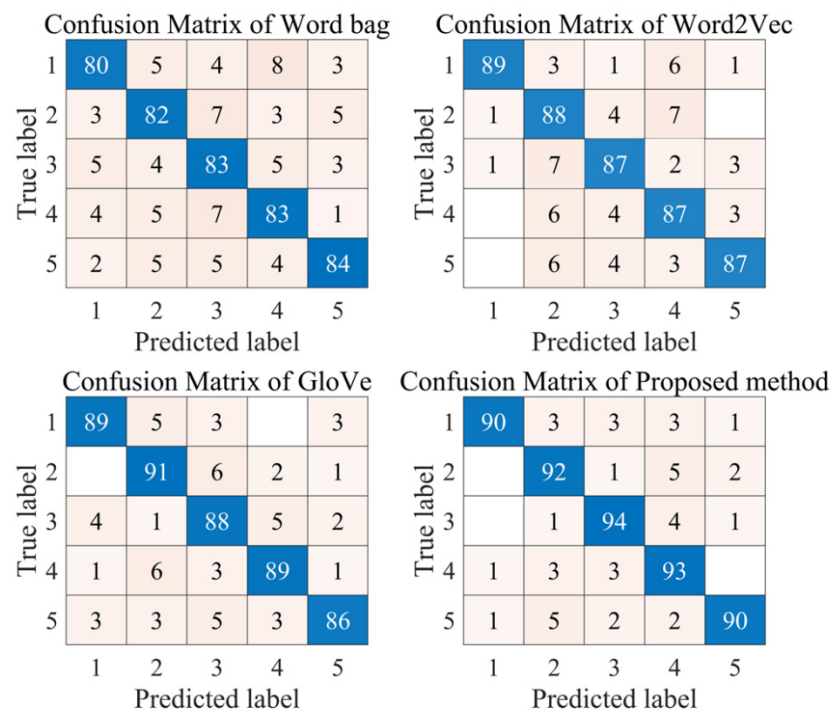


Figure 3. Confusion matrixes of different word embedding models.

4.3.2. Comparison of Various Ensemble Models

In order to prove that the MOGWO algorithm can effectively analyze and integrate the modeling capabilities of four base models, this paper compares the MOGWO algorithm with traditional Multi-Objective Particle Swarm Optimization (MOPSO) and NS-GAII algorithms to prove the practicability and advancement of the MOGWO algorithm. Table 4 shows the final recognition results obtained by different ensemble learning models. Figure 4 shows the curve of the variation of accuracy in the iterative process of different optimization algorithms. Based on Table 4 and Figure 4, the following conclusions can be drawn:

- (1) The ensemble learning model can effectively optimize the classification accuracy of a single model. This fully proves that ensemble learning is an effective method to improve the comprehensive recognition ability and generalization ability of models. The possible reason is that for different data, ensemble learning can dynamically optimize the weight according to the characteristics of different classifiers to establish the most suitable model for different data.
- (2) Compared with the traditional MOPSO algorithm and NSGAI algorithm, the MOGWO algorithm can achieve better results. This proves that the MOGWO algorithm has excellent weight optimization and decision-making ability in the field of integrated learning. The possible reason is that compared with the MOPSO algorithm and NSGAI algorithm, the MOGWO algorithm can obtain excellent optimization results with fewer parameters, which improves the overall convergence efficiency and global optimization ability of the model.

Table 4. Results of different ensemble models.

Model	Evaluation Metrics	Mood States					Average
		Great Satisfaction	Satisfaction	Neutral	Disappointment	Great Disappointment	
MOGWO	Accuracy	0.9760	0.9600	0.9700	0.9580	0.9720	0.9672
	Precision	0.9783	0.8846	0.9126	0.8692	0.9574	0.9204
	Recall	0.9000	0.9200	0.9400	0.9300	0.9000	0.9180
	F1	0.9375	0.9020	0.9261	0.8986	0.9278	0.9184
MOPSO	Accuracy	0.9600	0.9520	0.9460	0.9360	0.9660	0.9520
	Precision	0.9348	0.8800	0.8411	0.8091	0.9560	0.8842
	Recall	0.8600	0.8800	0.9000	0.8900	0.8700	0.8800
	F1	0.8958	0.8800	0.8696	0.8476	0.9110	0.8808
NSGAI	Accuracy	0.9600	0.9420	0.9500	0.9380	0.9620	0.9504
	Precision	0.9167	0.8318	0.8713	0.8350	0.9355	0.8780
	Recall	0.8800	0.8900	0.8800	0.8600	0.8700	0.8760
	F1	0.8980	0.8599	0.8756	0.8473	0.9016	0.8765

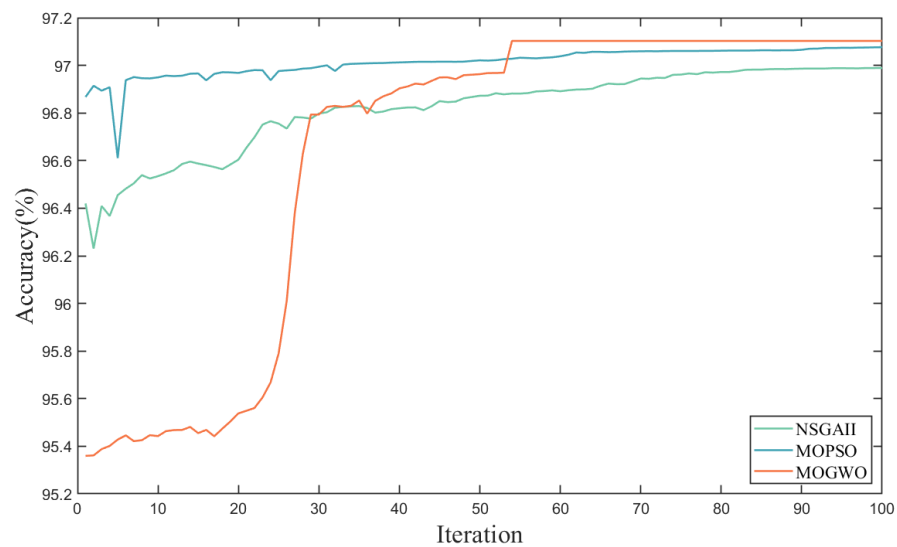


Figure 4. The iterative curve of different multi-objective optimization algorithms.

4.3.3. Comparison of Other Existing Models

In order to fully prove that the model proposed has excellent application value and is cutting edge, this paper compares the proposed model with five common benchmark models, including Seq2seq-Attention, Bilstm, RNN, TextCNN, and SVM. In these benchmark models, common text analysis models and some frontier models are included. Figure 5 shows the confusion matrix of the classification results of the proposed model. Figure 6 shows the ROC curves of the five emotions of the proposed model. Table 5 shows the AUC values of five emotions of the model proposed in this paper. Table 6 shows the emotion recognition results of all models. Based on Tables 5 and 6 and Figures 5 and 6, the following conclusions can be drawn:

- (1) Compared with traditional machine learning models, other deep learning models can achieve more excellent classification results. This fully proves that the deep learning algorithm has an excellent performance in the field of text emotion recognition. The possible reason is that the deep learning model effectively extracts the features of the original text through the multi-layer neural network structure and achieves more excellent results.
- (2) Compared with the traditional Bilstm neural network, the SEq2SEq_attention algorithm can obtain more accurate recognition results. This effectively proves that the attention mechanism effectively optimizes the performance of the neural network. The possible reason is that the attention mechanism can extract the deep correlation between the word vector data more deeply, which effectively improves the overall analysis and modeling ability of the model and achieves better recognition results.
- (3) The ensemble method proposed in this paper can achieve the best recognition results, which fully proves that the model proposed in this paper has excellent classification performance. The ensemble model adopted in this paper fully combines the advantages of each component. On the one hand, the feature representation of the original word vector is extracted by using the Word2Vec and GloVe word embedding models trained in this model, which effectively reduces the sparsity of the data. On the other hand, convolutional neural network and long and short-term memory networks were used for feature training, respectively, and four submodels were obtained by combining the attention mechanism. Finally, the integrated learning model based on a multi-objective gray Wolf optimizer can effectively make decisions on the weight of each sub-model and obtain the final emotion analysis results. Therefore, the model adopted in this paper can achieve excellent results in the field of emotion recognition.
- (4) Based on the results of AUC and ROC, it can be seen that the model proposed in this paper can obtain relatively accurate classification results for different emotion categories. In addition, the average AUC values for all classes of the model are satisfactory. Therefore, the model proposed in this paper has excellent emotional recognition ability and can stably identify different emotional categories.

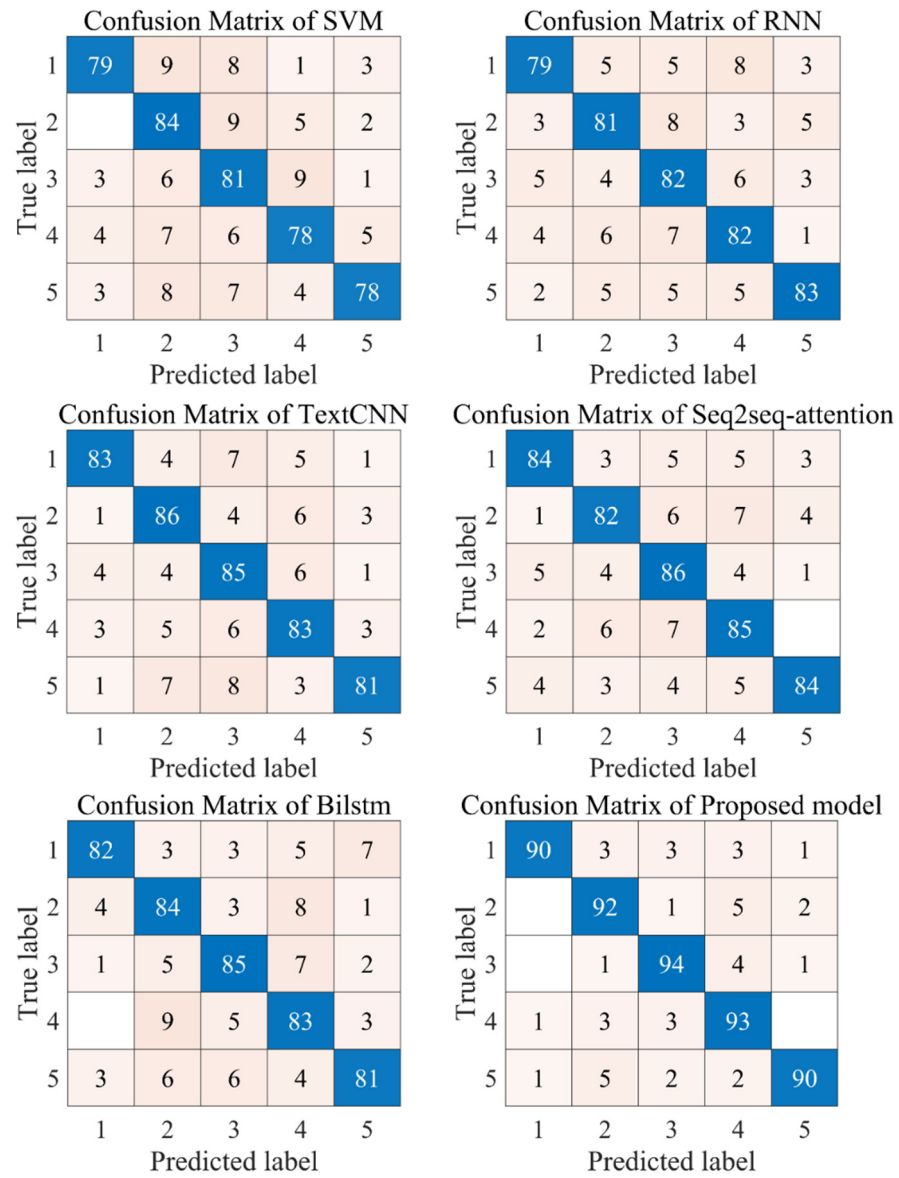


Figure 5. Confusion matrixes of the proposed model and other existing models.

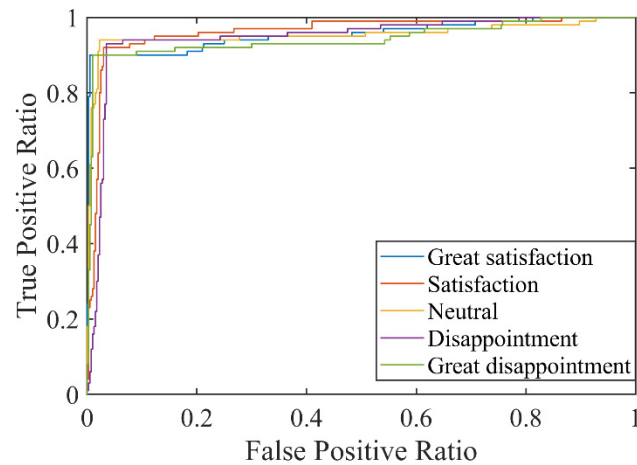


Figure 6. ROC curves of the proposed model.

Table 5. AUC values of five emotions of the proposed model.

Emotional Categories	Great Satisfaction	Satisfaction	Neutral	Disappointment	Great Disappointment	Average
AUC values	0.9546	0.9620	0.9541	0.9477	0.9429	0.9523

Table 6. Emotion recognition results of existing models.

Model	Evaluation Metrics	Mood States					Average
		Great Satisfaction	Satisfaction	Neutral	Disappointment	Great Disappointment	
The proposed model	Accuracy	0.9760	0.9600	0.9700	0.9580	0.9720	0.9672
	Precision	0.9783	0.8846	0.9126	0.8692	0.9574	0.9204
	Recall	0.9000	0.9200	0.9400	0.9300	0.9000	0.9180
	F1	0.9375	0.9020	0.9261	0.8986	0.9278	0.9184
Seq2seq-attention	Accuracy	0.9440	0.9320	0.9280	0.9280	0.9520	0.9368
	Precision	0.8750	0.8367	0.7963	0.8019	0.9130	0.8446
	Recall	0.8400	0.8200	0.8600	0.8500	0.8400	0.8420
	F1	0.8571	0.8283	0.8269	0.8252	0.8750	0.8425
TextCNN	Accuracy	0.9480	0.9320	0.9200	0.9260	0.9460	0.9344
	Precision	0.9022	0.8113	0.7727	0.8058	0.9101	0.8404
	Recall	0.8300	0.8600	0.8500	0.8300	0.8100	0.8360
	F1	0.8646	0.8350	0.8095	0.8177	0.8571	0.8368
Bilstm	Accuracy	0.9480	0.9220	0.9360	0.9180	0.9360	0.9320
	Precision	0.9111	0.7850	0.8333	0.7757	0.8617	0.8334
	Recall	0.8200	0.8400	0.8500	0.8300	0.8100	0.8300
	F1	0.8632	0.8116	0.8416	0.8019	0.8351	0.8307
SVM	Accuracy	0.9380	0.9080	0.9020	0.9180	0.9340	0.9200
	Precision	0.8876	0.7368	0.7297	0.8041	0.8764	0.8069
	Recall	0.7900	0.8400	0.8100	0.7800	0.7800	0.8000
	F1	0.8360	0.7850	0.7678	0.7919	0.8254	0.8012
RNN	Accuracy	0.9300	0.9220	0.9140	0.9200	0.9420	0.9256
	Precision	0.8495	0.8020	0.7664	0.7885	0.8737	0.8160
	Recall	0.7900	0.8100	0.8200	0.8200	0.8300	0.8140
	F1	0.8187	0.8060	0.7923	0.8039	0.8513	0.8144

5. Conclusions and Future Work

Online education and online course learning have become important learning methods for all classes after the rapid development of the epidemic. However, there are still many problems in online courses, and many teaching methods or the quality of courses cannot meet the requirements of learners. As key data reflecting the quality of online courses, sentiment recognition of user reviews is of great significance for evaluation and online courses. Emotional analysis of curriculum evaluation texts is helpful to identify effective curriculum suggestions, promote curriculum perfection and improvement, and is of great significance to improve the quality of online education.

Considering the complexity of network evaluation, the single-channel word vector may lead to information loss, and different deep network feature mining focuses are different, this paper proposes a two-channel deep network sentiment analysis model based on the MOGWO ensemble method, which abandons the sequential extraction of deep features by CNN and bidirectional LSTM. The information loss of sequence features in the convolution layer is avoided. The experimental results showed that this model has high accuracy and stability in the identification of emotional polarity and degree. On the other hand, this paper used Wikipedia's large corpus to pre-train Word2Vec and GloVe word embedding models to ensure that word vectors can obtain sufficient training information, which is of great significance for subsequent sentiment analysis.

The team's current research focuses on using the information contained in the text itself to identify the polarity and degree of emotion. As the main participants of the course, learners' personal information, such as course completion, personal assessment, and participation enthusiasm, also has varying degrees of correlation with the emotional tendency of the comment information. In the future, learners' information should be taken into consideration in sentiment analysis. In addition, after the completion of the sentiment analysis, the extraction and classification of key information of the evaluation text is also an important step to improve the quality of the course. The text classification problem is similar to the sentiment analysis, and relevant research can be carried out gradually in the future. In the future, the emotional recognition results of user evaluations can provide guidance for online courses, which can effectively improve the quality of online courses.

Author Contributions: Conceptualization, X.P., G.Y. and C.Y. (Chengqing Yu); methodology, X.P., G.Y. and X.M.; software, C.Y. (Chengqing Yu) and C.Y. (Chengming Yu); validation, X.P.; formal analysis, G.Y.; investigation, X.M.; resources, C.Y. (Chengqing Yu) and C.Y. (Chengming Yu); writing—original draft preparation, X.P., G.Y. and X.M.; writing—review and editing, C.Y. (Chengqing Yu) and C.Y. (Chengming Yu); visualization, X.P., G.Y. and X.M.; supervision, X.P.; funding acquisition, X.P., G.Y. and X.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The study did not report any data.

Acknowledgments: This study is fully supported by Science and education Joint project of Hunan Natural Science Foundation (2020JJ7031) and Fundamental Research Funds for the Central Universities (2019RC057).

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

References

- Jin, H.; Zhang, M.; He, Q.; Gu, J. Over 200 million students being taught online in China during COVID-19: Will online teaching become the routine model in medical education? *Asian J. Surg.* **2021**, *44*, 672. [[CrossRef](#)]
- Han, Z.-M.; Huang, C.-Q.; Yu, J.-H.; Tsai, C.-C. Identifying patterns of epistemic emotions with respect to interactions in massive online open courses using deep learning and social network analysis. *Comput. Hum. Behav.* **2021**, *122*, 106843. [[CrossRef](#)]
- Dong, C.; Cao, S.; Li, H. Young children's online learning during COVID-19 pandemic: Chinese parents' beliefs and attitudes. *Child. Youth Serv. Rev.* **2020**, *118*, 105440. [[CrossRef](#)] [[PubMed](#)]
- Cheng, P.; Ding, R. The effect of online review exercises on student course engagement and learning performance: A case study of an introductory financial accounting course at an international joint venture university. *J. Account. Educ.* **2021**, *54*, 100699. [[CrossRef](#)]
- Zhang, Q.; He, Y.-J.; Zhu, Y.-H.; Dai, M.-C.; Pan, M.-M.; Wu, J.-Q.; Zhang, X.; Gu, Y.-E.; Wang, F.-F.; Xu, X.-R.; et al. The evaluation of online course of Traditional Chinese Medicine for Medical Bachelor, Bachelor of Surgery international students during the COVID-19 epidemic period. *Integr. Med. Res.* **2020**, *9*, 100449. [[CrossRef](#)] [[PubMed](#)]
- Yang, X.; Xu, S.; Wu, H.; Bie, R. Sentiment Analysis of Weibo Comment Texts Based on Extended Vocabulary and Convolutional Neural Network. *Procedia Comput. Sci.* **2019**, *147*, 361–368. [[CrossRef](#)]
- Zhang, S.; Wei, Z.; Wang, Y.; Liao, T. Sentiment analysis of Chinese micro-blog text based on extended sentiment dictionary. *Future Gener. Comput. Syst.* **2018**, *81*, 395–403. [[CrossRef](#)]
- Aljuaid, H.; Iftikhar, R.; Ahmad, S.; Asif, M.; Tanvir Afzal, M. Important citation identification using sentiment analysis of in-text citations. *Telemat. Inform.* **2021**, *56*, 101492. [[CrossRef](#)]
- Widyassari, A.P.; Rustad, S.; Shidik, G.F.; Noersasongko, E.; Syukur, A.; Affandy, A.; Setiadi, D.R.I.M. Review of automatic text summarization techniques & methods. *J. King Saud Univ.—Comput. Inf. Sci.* **2020**, in press. [[CrossRef](#)]
- Ghulam, H.; Zeng, F.; Li, W.; Xiao, Y. Deep Learning-Based Sentiment Analysis for Roman Urdu Text. *Procedia Comput. Sci.* **2019**, *147*, 131–135. [[CrossRef](#)]
- Do, H.H.; Prasad, P.W.C.; Maag, A.; Alsadoon, A. Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. *Expert Syst. Appl.* **2019**, *118*, 272–299. [[CrossRef](#)]
- Kumhar, S.H.; Kirmani, M.M.; Sheetlani, J.; Hassan, M. Word Embedding Generation for Urdu Language using Word2vec model. *Mater. Today Proc.* **2021**, in press. [[CrossRef](#)]

13. Liu, R.; Sisman, B.; Lin, Y.; Li, H. FastTalker: A neural text-to-speech architecture with shallow and group autoregression. *Neural Netw.* **2021**, *141*, 306–314. [[CrossRef](#)]
14. Sakketou, F.; Ampazis, N. A constrained optimization algorithm for learning GloVe embeddings with semantic lexicons. *Knowl. Based Syst.* **2020**, *195*, 105628. [[CrossRef](#)]
15. Sharma, A.K.; Chaurasia, S.; Srivastava, D.K. Sentimental Short Sentences Classification by Using CNN Deep Learning Model with Fine Tuned Word2Vec. *Procedia Comput. Sci.* **2020**, *167*, 1139–1147. [[CrossRef](#)]
16. Mirończuk, M.M.; Protasiewicz, J. A recent overview of the state-of-the-art elements of text classification. *Expert Syst. Appl.* **2018**, *106*, 36–54. [[CrossRef](#)]
17. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [[CrossRef](#)] [[PubMed](#)]
18. Mikolov, T.; Corrado, G.; Kai, C.; Dean, J. Efficient Estimation of Word Representations in Vector Space. *arXiv* **2013**, arXiv:1301.3781.
19. Naderalvojud, B.; Sezer, E.A. Sentiment aware word embeddings using refinement and senti-contextualized learning approach. *Neurocomputing* **2020**, *405*, 149–160. [[CrossRef](#)]
20. Onan, A. Two-Stage Topic Extraction Model for Bibliometric Data Analysis Based on Word Embeddings and Clustering. *IEEE Access* **2019**, *7*, 145614–145633. [[CrossRef](#)]
21. Muhammad, P.F.; Kusumaningrum, R.; Wibowo, A. Sentiment Analysis Using Word2vec And Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews. *Procedia Comput. Sci.* **2021**, *179*, 728–735. [[CrossRef](#)]
22. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
23. Cao, X.; Li, J.; Wang, R.; Wang, Y.; Niu, Q.; Shi, J. Calibrating GloVe model on the principle of Zipf’s law. *Pattern Recognit. Lett.* **2019**, *125*, 715–720. [[CrossRef](#)]
24. Kamkarhaghighi, M.; Makrehchi, M. Content Tree Word Embedding for document representation. *Expert Syst. Appl.* **2017**, *90*, 241–249. [[CrossRef](#)]
25. Li, S.; Pan, R.; Luo, H.; Liu, X.; Zhao, G. Adaptive cross-contextual word embedding for word polysemy with unsupervised topic modeling. *Knowl. Based Syst.* **2021**, *218*, 106827. [[CrossRef](#)]
26. Nozza, D.; Manchanda, P.; Fersini, E.; Palmonari, M.; Messina, E. LearningToAdapt with word embeddings: Domain adaptation of Named Entity Recognition systems. *Inf. Process. Manag.* **2021**, *58*, 102537. [[CrossRef](#)]
27. Khan, J.; Lee, Y.K. LeSSA: A Unified Framework based on Lexicons and Semi-Supervised Learning Approaches for Textual Sentiment Classification. *Appl. Sci.* **2019**, *9*, 5562. [[CrossRef](#)]
28. Wunderlich, F.; Memmert, D. Innovative Approaches in Sports Science-Lexicon-Based Sentiment Analysis as a Tool to Analyze Sports-Related Twitter Communication. *Appl. Sci.* **2020**, *10*, 431. [[CrossRef](#)]
29. Mäntylä, M.V.; Graziotin, D.; Kuuttila, M. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Comput. Sci. Rev.* **2018**, *27*, 16–32. [[CrossRef](#)]
30. Chintalapudi, N.; Battineni, G.; Canio, M.D.; Sagaro, G.G.; Amenta, F. Text mining with sentiment analysis on seafarers’ medical documents. *Int. J. Inf. Manag. Data Insights* **2021**, *1*, 100005. [[CrossRef](#)]
31. Song, C.; Wang, X.-K.; Cheng, P.-F.; Wang, J.-Q.; Li, L. SACPC: A framework based on probabilistic linguistic terms for short text sentiment analysis. *Knowl. Based Syst.* **2020**, *194*, 105572. [[CrossRef](#)]
32. Jia, K. Chinese sentiment classification based on Word2vec and vector arithmetic in human–robot conversation. *Comput. Electr. Eng.* **2021**, *95*, 107423. [[CrossRef](#)]
33. Tran, T.K.; Phan, T.T. Deep Learning Application to Ensemble Learning-The Simple, but Effective, Approach to Sentiment Classifying. *Appl. Sci.* **2019**, *9*, 2760. [[CrossRef](#)]
34. Han, K.X.; Chien, W.; Chiu, C.C.; Cheng, Y.T. Application of Support Vector Machine (SVM) in the Sentiment Analysis of Twitter DataSet. *Appl. Sci.* **2020**, *10*, 1125. [[CrossRef](#)]
35. Nassif, A.B.; Elnagar, A.; Shahin, I.; Henno, S. Deep learning for Arabic subjective sentiment analysis: Challenges and research opportunities. *Appl. Soft Comput.* **2021**, *98*, 106836. [[CrossRef](#)]
36. Liao, W.; Zeng, B.; Liu, J.; Wei, P.; Cheng, X.; Zhang, W. Multi-level graph neural network for text sentiment analysis. *Comput. Electr. Eng.* **2021**, *92*, 107096. [[CrossRef](#)]
37. Ullah, M.A.; Marium, S.M.; Begum, S.A.; Dipa, N.S. An algorithm and method for sentiment analysis using the text and emoticon. *ICT Express* **2020**, *6*, 357–360. [[CrossRef](#)]
38. Shi, L.; Jianping, C.; Jie, X. Prospecting Information Extraction by Text Mining Based on Convolutional Neural Networks—A Case Study of the Lala Copper Deposit, China. *IEEE Access* **2018**, *6*, 52286–52297. [[CrossRef](#)]
39. Dong, S.; Yu, C.; Yan, G.; Zhu, J.; Hu, H. A Novel ensemble reinforcement learning gated recursive network for traffic speed forecasting. In Proceedings of the 2021 Workshop on Algorithm and Big Data, Fuzhou, China, 12–14 March 2021; pp. 55–60.
40. Xia, Y.; Chen, K.; Yang, Y. Multi-label classification with weighted classifier selection and stacked ensemble. *Inf. Sci.* **2021**, *557*, 421–442. [[CrossRef](#)]
41. Nnabuife, S.G.; Kuang, B.; Whidborne, J.F.; Rana, Z. Non-intrusive classification of gas-liquid flow regimes in an S-shaped pipeline riser using a Doppler ultrasonic sensor and deep neural networks. *Chem. Eng. J.* **2021**, *403*, 126401. [[CrossRef](#)]

42. Minakova, S.; Stefanov, T. Buffer sizes reduction for memory-efficient CNN inference on mobile and embedded devices. In Proceedings of the 2020 23rd Euromicro Conference on Digital System Design (DSD 2020), Kranj, Slovenia, 26–28 August 2020; pp. 133–140.
43. Yan, G.; Yu, C.; Bai, Y. Wind Turbine Bearing Temperature Forecasting Using a New Data-Driven Ensemble Approach. *Machines* **2021**, *9*, 248. [[CrossRef](#)]
44. Zhao, J.F.; Mao, X.; Chen, L.J. Learning deep features to recognise speech emotion using merged deep CNN. *IET Signal Process.* **2018**, *12*, 713–721. [[CrossRef](#)]
45. Cai, H.J.; Chen, T. Multi-dimension CNN for hyperspectral image classificaton. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1275–1278.
46. Liu, X.; Qin, M.; He, Y.; Mi, X.; Yu, C. A new multi-data-driven spatiotemporal PM2. 5 forecasting model based on an ensemble graph reinforcement learning convolutional network. *Atmos. Pollut. Res.* **2021**, *12*, 101197. [[CrossRef](#)]
47. Yang, G.; Xu, H.Z. A Residual BiLSTM Model for Named Entity Recognition. *IEEE Access* **2020**, *8*, 227710–227718. [[CrossRef](#)]
48. Liao, F.; Ma, L.L.; Yang, D.J. Research on Construction Method of Knowledge Graph of US Military Equipment Based on BiLSTM model. In Proceedings of the 2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), Shenzhen, China, 9–11 May 2019; pp. 146–150.
49. Ma, W.; Yu, H.Z.; Zhao, K.; Zhao, D.S.; Yang, J.; Ma, J. Tibetan location name recognition based on BiLSTM-CRF model. In Proceedings of the 2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CYBERC), Guilin, China, 17–19 October 2019; pp. 412–416.
50. Zhang, P.F.; Li, F.H.; Du, L.D.; Zhao, R.J.; Chen, X.X.; Yang, T.; Fang, Z. Psychological Stress Detection According to ECG Using a Deep Learning Model with Attention Mechanism. *Appl. Sci.* **2021**, *11*, 2848. [[CrossRef](#)]
51. Chen, W.J.; Li, J.L. Forecasting Teleconsultation Demand Using an Ensemble CNN Attention-Based BiLSTM Model with Additional Variables. *Healthcare* **2021**, *9*, 992. [[CrossRef](#)] [[PubMed](#)]
52. Du, J.; Cheng, Y.Y.; Zhou, Q.A.; Zhang, J.M.; Zhang, X.Y.; Li, G.; IOP. Power load forecasting using BiLSTM-attention. In Proceedings of the 2019 5th International Conference on Environmental Science and Material Application, Xi'an, China, 15–16 December 2019.
53. Zapotecas-Martínez, S.; García-Nájera, A.; López-Jaimes, A. Multi-objective grey wolf optimizer based on decomposition. *Expert Syst. Appl.* **2019**, *120*, 357–371. [[CrossRef](#)]
54. Gupta, S.; Deep, K. Enhanced leadership-inspired grey wolf optimizer for global optimization problems. *Eng. Comput.* **2020**, *36*, 1777–1800. [[CrossRef](#)]
55. Ling, O.Y.; Theng, L.B.; Chai, A.; McCarthy, C. A model for automatic recognition of vertical texts in natural scene images. In Proceedings of the 2018 8th IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2018), Penang, Malaysia, 23–25 November 2018; pp. 170–175.
56. Liu, M.F.; Xie, Z.C.; Huang, Y.X.; Jin, L.W.; Zhou, W.Y. Distilling GRU with data augmentation for unconstrained handwritten text recognition. In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; pp. 56–61.
57. Nguyen, H.T.; Nguyen, C.T.; Nakagawa, M. ICFHR 2018-competition on Vietnamese online handwritten text recognition using HANDS-VNOnDB (VOHTR2018). In Proceedings of the 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), Niagara Falls, NY, USA, 5–8 August 2018; pp. 494–499.
58. Qin, Y.; Zhang, Z. Summary of scene text detection and recognition. In Proceedings of the 15th IEEE Conference on Industrial Electronics and Applications (ICIEA 2020), Kristiansand, Norway, 9–13 November 2020; pp. 85–89.