Emotion Analysis Model of MOOC Course Review Based on BiLSTM

Shen Ji (✉), Tan Fangbi
Zhongnan University of Economics and Law, Wuhan, China
`jishen012@126.com`

Abstract—Online course reviews can objectively reflect the emotional tendency of learners towards the learning effect. This paper proposes a deep neural network-based sentiment analysis model for MOOC course reviews. The model uses the Bidirectional Long Short-Term Memory Network (BiLSTM) to analyze Chinese semantic. To deal with the imbalance of the training data set, this paper introduces two methods to balance it and add a dropout mechanism to prevent the overfitting model. The model is then applied to the emotional evaluation of the MOOC course of "Fundamentals of College Computer Application." The application results show that the model has achieved good accuracy and can well realize the emotional orientation analysis of online course reviews to provide a valuable reference for Course Builders.

1   Introduction

With the rapid development of Internet technology, MOOC's online courses have attracted more and more attention from educators. Especially in the early novel coronavirus pneumonia epidemic spread in early 2020, more and more colleges and universities use online courses like MOOC to complete undergraduate teaching in order to avoid the risk of offline centralized classroom teaching. However, the construction cycle of online courses in Colleges and Universities in China is generally short. At present, the quality of online courses represented by MOOC is uneven. How to accurately evaluate the course to optimize and complete the course content, and the teaching modes and methods are fundamental. In learning, MOOC learners will produce a variety of learning behavior data, which truly reflects the state of learners in learning. Therefore, appropriate data mining and analysis technology to analyze learning behavior data can find potential and valuable information and provide data support for curriculum evaluation [1-2]. In the past research on curriculum learning behavior analysis, more attentions have been paid to the structured curriculum behavior data,

such as the number of course selection and change trend, the length and times that learners watch videos of each chapter, the number of participation in discussions and the examination results. For unstructured interactive text data, especially learners' comments on the course, are paid less attention. Online course review is an important feedback of learners' perception of course learning. By analyzing the course review's emotional polarity, we can quickly understand the learners' views on the content. On the one hand, it can provide personalized intervention according to learners' emotional state in order to provide a personalized learning experience with humanistic care for learners; on the other hand, it can comment on the emotional distribution according to the overall curriculum so that the overall online teaching level and effect of the course are evaluated. The existing problems and shortcomings are analyzed, and the policy and strategy of course reform are established.

Emotions are people's views or emotions that assess their attitudes towards products, services, or organizations. Sentiment analysis is a task in natural language processing called tendency analysis, opinion extraction, opinion mining, emotion mining, and subjective analysis. It can process, analyze, summarize, and infer the subjective text with emotion [3]. Generally, the effective analysis's main purpose is to mine bipolar attitudes or views from speakers on a certain topic (or topics). This view or attitude may be the personal assessment or judgment of the speaker or his (her) emotional state or intention to communicate with others [4]. The sentiment analysis of MOOC course review in this paper belongs to text orientation analysis; that is, judge the emotional category (positive or negative) expressed by each course review text. To accomplish this task, two key problems need to be solved. The first is the extraction of emotional elements and the analysis of emotional polarity. Given this issue, the deep learning method based on Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) has been widely used in image processing and natural language processing [5]. RNN has made great achievements in the application of textual sentiment analysis. Schuster and Paliwal first proposed a Bidirectional Recurrent Neural Network (BRNN) in 1997. This network can focus on the context of the sentence simultaneously to use more information for semantic analysis and prediction. Structurally, BRNN is composed of two RNNs with opposite directions. These two RNNs are connected to the same output layer to receive the hidden layer information of the previous time and the next moment as the input and realize the purpose of paying attention to the context simultaneously [6]. Song et al. designed a bidirectional text generation model by using RNN [7]. The model can use two RNNs to construct context statements according to a given the word and introduce a position vector into the model to solve the problem of location information loss of specified words. Al-Smadi et al. applied the Long Short-Term Memory network (LSTM) to the emotional analysis of Arab Hotel Reviews. LSTM networks with an element and character level were constructed to complete emotion analysis tasks of different granularity and coupled them. The application results show that the coupled model's classification effect is significantly better than that of the emotion analysis model with two granularities [8]. The second problem is the imbalance of MOOC review data. No matter what method is used to extract emotional elements and polarity analysis of MOOC review data, the review data set is an important reference element in model construction. However,

MOOC reviews' data is obviously unbalanced, and the course reviews tend to distribute to a curtain pole. For example, a course review is generally better, and the number of positive reviews accounts for the vast majority of all reviews. The training based on unbalanced data will lead to emotion classification seriously biased towards the emotion category with more samples and ignore the emotion category with fewer sample numbers. Thus, the classification performance is greatly reduced [9]. Therefore, it is necessary to balance the data set to solve this problem at present. The research on data balance of text data sets mainly focuses on data preprocessing, feature extraction, and classification model algorithm. The balance methods in data preprocessing mainly use resampling methods, such as oversampling and undersampling. These methods are easy to implement and achieve good results in some classification tasks. Thus they are widely used. However, the MOOC review data in this paper belongs to the text data set, which is difficult to obtain, and the data collective quantity is limited. Therefore, it is difficult to use the resampling method to realize it.

Many research results have shown that the deep learning method's text sentiment analysis model has a better application effect than the traditional emotion dictionary and machine learning method [10-11]. However, the application of deep text classification methods such as LSTM model to sentiment analysis of MOOC online review data is relatively less, and its application effect needs to be further verified. On the other hand, MOOCs online reviews generally have obvious emotional polarity bias, and the positive and negative emotional data are often seriously unbalanced. However, traditional data balancing processing has limitations on MOOC reviews' application effect as text data. Therefore, this paper will use the deep learning method to construct the MOOC course review's emotion analysis model. Thus, two text data generation methods are introduced to deal with the MOOC review data set's imbalance. Furthermore, the online review data of the "Fundamentals of College Computer Application "MOOC course is combined to verify its effect in evaluating emotional orientation, of course. These application results of the two-text data balance processing methods are compared and analyzed to find the effective means of personalized auxiliary teaching and course evaluation of MOOC online course

2 Emotional Analysis Model of MOOC Course Review

When the deep learning method is used for text analysis, the cyclic neural network can store the sequence information persistently. However, the traditional recurrent neural network still has the problem of long-term dependence [12]. Hochreiter and Schmidhuber proposed Long Short-Term Memory Network (LSTM) in 1997[13]. It adds a memory unit to each neural unit of the hidden layer based on the common neural network. The hidden layer controls memory or forgetting some information through four controllable gates. LSTM can effectively solve the long-term dependency problem of ordinary RNN and has been widely used in natural language processing tasks [14-15]. However, the simple LSTM model can only encode the text in one direction. The model can only analyze the information above at the current time but not obtain the association information below it. In order to solve this problem, researchers

proposed a Bi-directional Long Short-Term Memory network (BiLSTM) model based on BRNN [14]. Therefore, the model consists of a forward LSTM and a reverse LSTM. The forward LSTM is input by the first word at the beginning of the text, while the reverse LSTM is input from the last word of the text to capture the two-way semantic dependency better. Meanwhile, the gradient explosion and gradient disappearance problems in the traditional RNN can be solved. Thus, this paper constructs a sentiment analysis model of MOOC course review based on BiLSTM, and its structure is shown in Figure 1.
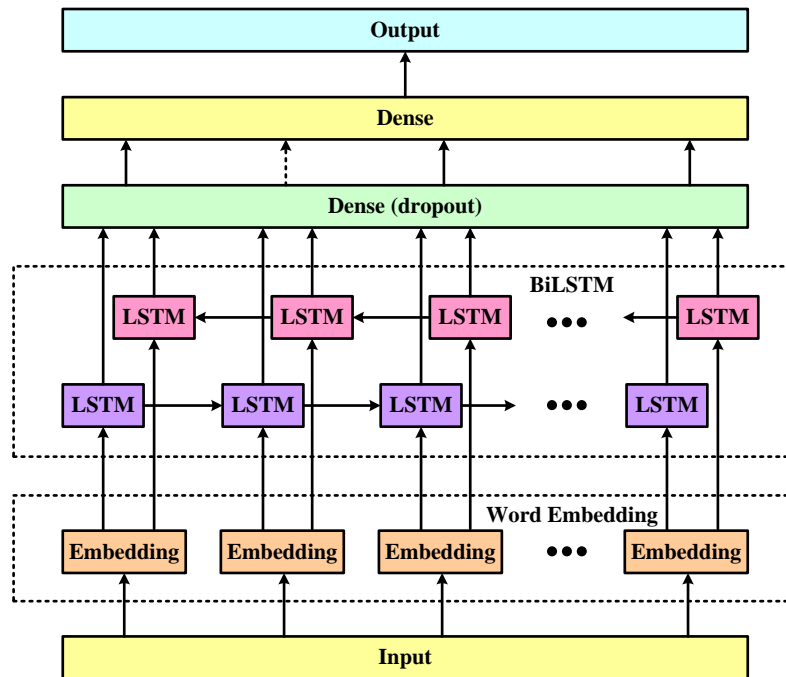

Fig. 1. Emotional analysis model of MOOC course review

The contents of each layer in the model structure are as follows:

1. Input layer: Taking a MOOC course review text as the input, a dictionary is created according to the word segmentation results. Each word in the review is coded, and the coding sequence is used as the input data of the model.
2. Word embedding layer: The word embedding layer uses the N-gram model to transform each word in the text coding sequence into the corresponding low dimensional word vector to solve the data sparsity problem in one hot coding. On the other hand, the vectorization of words can better reflect the relevance between words, thus further ensuring the model's training effect.

3. BiLSTM layer: The word vector sequence is input into positive LSTM and negative LSTM, respectively, in a different order, so that the model can extract the emotional features of the text more fully with the context semantics.
4. Dropout mechanism: In order to avoid over-fitting problems in the model, a dropout mechanism is added [15]. Its discarding probability is set to 0.5; that is, half of the neurons are inactivated. The overall complexity of the model is reduced.
5. Full connectivity layer: A full connection layer with 128 output dimension is constructed, and the activation function is Relu to avoid gradient explosion and gradient disappearance in the process of backpropagation.
6. Output layer: Considering that the model is essentially a dichotomous problem, the final output is two types of emotion (positive/negative). A full connection layer with an output dimension of two is added at the end of the model, and the activation function is Softmax; that is, the probability distribution of two emotional types is output. The type with high probability is selected as the result of the emotional judgment.

3   Text Data Balance Processing Method

To realize the balanced processing of MOOC comment text data, this paper introduces the simple copy method and Easy Data Augmentation Techniques (EDA). In the simple replication method, the data balance is achieved by extending the replication of negative reviews. The principle of this method is simple and easy to implement. However, it will generate many duplicate data samples, which are completely consistent in semantic features and cannot provide more feature reference information for the classifier in model training and learning. Thus, the application effect may be minimal. The text data enhancement algorithm is generally used to improve the model's anti-interference ability in the text generation task. For example, in the machine translation task, noise-adding statements similar to the source statement are created by replacing and deleting words. The main contents of the text data enhancement algorithm include (1) Synonym replacements. Several words are randomly selected from the text, and then the synonyms are randomly extracted from the synonym dictionary and replaced. (2) Random insertion. A word is randomly selected from the text and then randomly selected from the synonym set to insert into the original sentence's random position. The process can be repeated many times. (3) Random exchange. Two words are randomly selected for position exchange, and the process can be repeated many times. (4) Random deletion. The deletion factor of each word is calculated. If the deletion factor is greater than a certain threshold, the word will be deleted. In this paper, the above method will be used to process a polarity of the course review text and supplement the number of negative comments by generating synonymous evaluation statements.

4 Experiments and Results

To verify the effect of the model, this paper selects the course "Fundamentals of College Computer Application" offered by Zhongnan University of Finance and Law in the MOOC platform as the application object. The emotional analysis experiment is carried out on the students' reviews to obtain information on learners' emotional tendency to the teaching effect. The overall design idea of the experiment is shown in Figure 2.
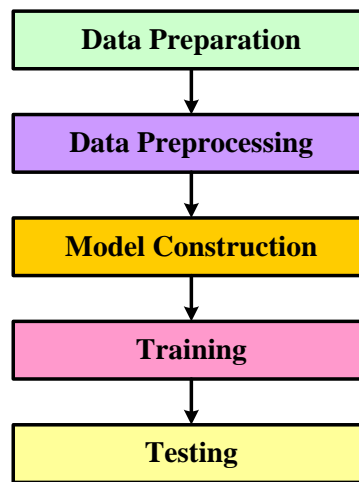
```
┌─────────────────────────┐
│    Data Preparation     │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│    Data Preprocessing   │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│    Model Construction   │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│        Training         │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│         Testing         │
└─────────────────────────┘
```

Fig. 2. Experimental design

4.1 Data preparation

To obtain the review data of MOOCS for sentiment analysis, this paper uses the request Library in Python language to build a web crawler. Enter the course review page of "Fundamentals of College Computer Application" of Zhongnan University of Finance and law under the MOOC platform to crawl the students' reviews and save them in the XLSX file. A total of 1657 course review data are obtained. Considering that the amount of data may not meet the needs of model parameter training, the text continues to crawl through the MOOC platform of Chinese universities to obtain 5037-course reviews of "computer foundation" courses offered by other universities.

4.2 Data pre-processing

1. Data clean: Some course reviews may be invalid without practical significance. This kind of text belongs to data noise in this experiment. It is necessary to clean them up to avoid the impact on feature extraction. Firstly, according to the reviews'

length, too short (less than 3 words) ones are eliminated, and then manually clean the content of each review. Finally, 4981 meaningful course review data is obtained.

2. Data annotation: Considering that the amount of data involved in the experiment is not large, artificial emotion annotation is used for all review data. Positive reviews are marked as 1, and negative reviews are marked as 0. A total of positive 4097 and negative 884 data are obtained.

3. Data balance: A simple analysis of the data, positive reviews accounted for more than 80% of the total data set, which belongs to unbalanced data. To solve this problem, in this paper, two methods of simple copy and text enhancement are used to balance the comments, and two methods obtain two sets of experimental data sets. Each data set contains 8000 test/training data, including 4097 positive reviews and 3903 negative reviews. Each group of data will be divided into a training set and test set in the ratio of 8:2.

4. Text segmentation: Word segmentation is an important work in the early stage of data processing. The jieba Chinese word segmentation library is used to segment each review text data to remove the characters without emotional information such as numbers, punctuation marks (except exclamation marks and question marks), and stop words. According to the word segmentation results, a dictionary is created to encode each review data as the sentiment analysis model's input.

4.3 Model building

According to the model network structure in Figure 1, the text uses Baidu's paddle deep learning framework to complete the construction of the MOOC course reviews sentiment analysis model. The main model parameter settings are shown in Table 1. Considering the small amount of training and test data, this paper selects a small batch size (64). To ensure the model to achieve convergence in the training process, the overall number of iterations is set to 15. In terms of the loss function, the cross-entropy loss function commonly used in the classification model is selected, and Adam optimizer is selected to accelerate the model's convergence rate.

Table 1. Main model parameters setting

Parameter	Value
Word vector dimension (word embedding size)	128
BiLSTM Hidden dimension (hidden size)	128
Dropout probability	0.5
Full connectivity dimension	128
Training/testing batch size	64
Training round (epoch)	15
Loss function	Binary_crossentropy
Optimizer	Adam
Learning rate	0.0002

4.4 Model training and test results

After the model construction is completed, the data sets obtained by simple replication method and EDA method are imported for training and testing. Model Accuracy is used to describe the training accuracy, precision, recall, and F1 score models to describe the model's test accuracy. After 15 rounds of training, the training accuracy loss curve obtained by these two data balance methods is shown in Fig. 3. It can be seen from Figure 3 that the model accuracy loss curves obtained by using the data sets from these two data balancing methods as input are very similar. The model converges at about 12 iterations and achieves high training accuracy (all above 90%). Among them, the model's training results obtained by the simple replication method are better, and the highest training accuracy is 97% with relatively fast convergence speed. While the highest training accuracy of the model obtained by the EDA method is 95%.

After the parameter training, the test data set is used to evaluate the effect of model emotion analysis, and the model test results are shown in Table 2. It can be seen from Table 2 that the test results of the model constructed with the raw data set without processing are relatively poor, and the accuracy and recall rate are both around 0.8. It can be seen that the unbalanced original data set does have a significant impact on the results of emotional feature extraction and classification. Additionally, the model test results obtained using different data balance methods are similar to the training results. The overall test effect of the model is good, most of the precision parameters of the model test are above 0.9, and the recall rate of the EDA method is less than 0.9 (0.89). Compared with the training accuracy of the models, there is no obvious overfitting. Similar to the model training's training results, the model accuracy, recall rate, and F1 score are better when the model test results obtained by using different data balance methods are used as input.
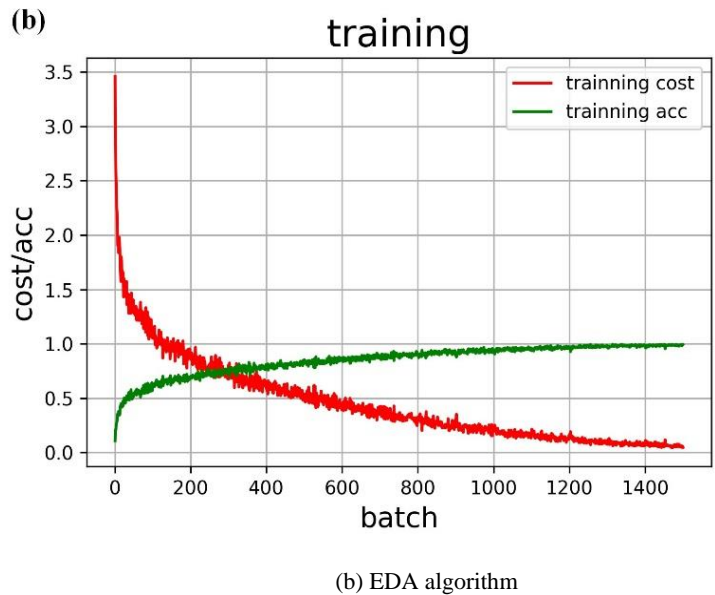
(a)
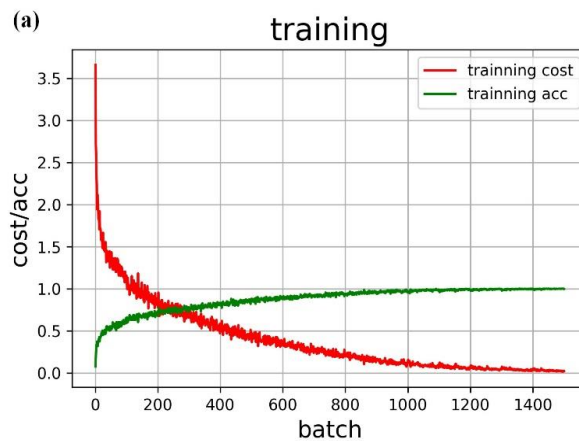
(a) simple copy method

(b)

(b) EDA algorithm

Fig. 3. Model training accuracy loss curve

Table 2. Model test results

Data balancing method	Accuracy	Recall	F1 score
No.	0.79	0.83	0.81
Simple replication	0.92	0.96	0.94
EDA	0.91	0.89	0.90
Improved EDA Method	0.94	0.97	0.96

To further analyse the impact of unbalanced data sets and various balanced processing methods on the results of emotional analysis, the data set balanced by each method is used as input, and the confusion matrix of the model test is shown in Table 3. It can be seen that when the data set is not balanced, the prediction effect of the model on the emotional polarity is relatively poor, especially for the prediction of negative emotional comments. The model without positive comments predicts a considerable number of negative comment texts. This may be because most positive comments in the training set account for the majority of the training set, which leads to learning too many positive emotional features in the training process. The model has insufficient recognition ability for negative emotion sentences. Furthermore, it is noted that when using the simple replication method, the model's prediction effect on positive emotional tendency is better than that on the negative emotional tendency. Compared with the simple replication method, when using the EDA method, the prediction effect of positive and negative data is relatively balanced, and the difference between accuracy and recall rate is small. However, the overall prediction effect is not

as good as the former. When using the EDA method, the model mispredicted more positive reviews than negative ones. This may be that the feature range of negative reviews is enlarged when using EDA Method to generate synonyms of negative reviews. As a result, the model cannot distinguish positive comments from negative comments. Generally, the model test's effect is better than that of non-balanced processing after using the simple copy method and EDA method to balance the comment text. It is consistent with the results obtained by Wei and Zhou and further confirms the importance of data balance processing in text sentiment classification tasks [18].

Table 3. Confusion matrix of each model test

		Predicted value (positive)	**Predicted value (negative)**
No.	Predicted value (positive)	3420	677
	Predicted value (negative)	906	2997
Simple replication	Predicted value (positive)	3943	154
	Predicted value (negative)	346	3557
EDA Method	Predicted value (positive)	3657	440
	Predicted value (negative)	372	3531
Improved EDA Method	Predicted value (positive)	3971	126
	Predicted value (negative)	239	3704

4.5 Result analysis

From the results of Fig.3 and Table 2, it is known that the accuracy of the emotion analysis model of the MOOC course review of "Foundation of College Computer Application "proposed in this paper is more than 90%. Therefore, it can be considered that using BiLSTM to construct a deep neural network can reasonably analyze the emotion of MOOC course review. Besides, due to the obvious imbalance of the experimental data set, this paper uses two data methods to balance it. The EDA method should have better prediction accuracy than the simple copy method because of the text enhancement algorithm. However, by comparing the results of these two methods, it can be seen that the accuracy of the model obtained by using the EDA Method to enhance the negative data set text is not as accurate as of that of the simple copy method. Although it is easy to synthesize sentences or even texts by using the text generation tools in the EDA method, the reason may be that it is easy to synthesize sentences or even texts. Actually, the so-called "generated text" is not valid in most cases. These "generated texts" enhance the text to a certain extent, but they also interfere with the original model by randomly adding and deleting words. In some cases, the original meaning of the sentence is completely changed. For example, the random deletion rule may remove the key emotional words in a review. This may make the lead type learn completely different feature information in the training process, which greatly impacts the actual application effect of the model. Wei and Zhou also put forward similar conclusions. They pointed out that the EDA algorithm is mainly used to provide noise samples in text generation tasks to enhance the model's generalization performance. For classification tasks, excessive data enhancement processing has a

limited improvement on the model and even leads to the decline of the model effect. For example, the random insertion operation in the EDA algorithm may lead to the loss of semantic structure and order of original training data, and the emotional keywords of sentences are not considered in the selection process of insertion words. It leads to the insertion of new words that often do not contain valuable information, and the effect is limited in the diversity of data expansion. Although the random method in EDA can take care of all the words, it lacks the weight allocation of keywords. If the random operation happens to be the words with the strongest emotional characteristics, it may lead to the complete change of semantic information and even emotional polarity. In contrast, the synonym replacement in the EDA algorithm is more in line with the processing requirements of data balancing and does not cause too many negative effects. Therefore, this paper improves the EDA algorithm, eliminates the random enhancement operation in the algorithm, and only retains the synonym replacement. The model test results are shown in Table 2 and Table 3. It can be seen that the model effect of using the improved EDA algorithm for balancing processing is relatively the best among all methods, and the overall prediction accuracy has been significantly improved [18].

5   Conclusion

According to the needs of the current online course quality evaluation and Optimization Reform in Colleges and Universities, this paper proposes a sentiment analysis model of MOOC course evaluation based on BiLSTM. It is applied to the emotional analysis of MOOC course evaluation of "Fundamentals of College Computer Application. " The results show that the model can be used to analyse the emotional tendency of curriculum rating. Besides, to solve the problem of data imbalance, the simple copy method results are better than those by the EDA method. This paper analyses the reasons for this situation, proposes an improved EDA Method for data processing, and further improves emotional polarity's prediction accuracy. This paper's results show that the deep neural network can effectively solve the problem of emotional orientation judgment of online course reviews. The solution to data imbalance of online course review is proposed. The text's relevant results can be applied to the emotional analysis tasks of various online course reviews and make a systematic analysis of the distribution of emotional polarity of curriculum evaluation. It provides strong data support for the construction and improvement of the curriculum. On the other hand, there are some deficiencies in the research and design process. In the process of emotion analysis modelling, the model is simplified to reduce complexity. For example, all the review data sets are simply divided into positive and negative ones. Without considering the neutral course reviews between them, it may have a certain impact on the model's actual application effect. At the same time, in the process of model application, it is found that some course reviews may be composed of multiple short sentences, and each sentence may contain an emotional transition. This is not fully considered in the current model. In future research, we will fully consider the above two problems in the modelling process and optimize and improve the EDA method. Thus,

it can better solve the imbalance problem of the review data set in the emotional analysis of curriculum reviews and further improve the model's practical application effect.

6   Acknowledgment

7   References

[1] Papadakis, S., Kalogiannakis, M., Sifaki, E., & Vidakis, N. (2017). Access moodle using smart mobile phones. A case study in a Greek University. In Interactivity, Game Creation, Design, Learning, and Innovation. Springer, Cham 376-385. https://doi.org/10.1007/978-3-319-76908-0_36

[2] Papadakis, S., Kalogiannakis, M., Sifaki, E., & Vidakis, N. (2018). Evaluating Moodle use via Smart Mobile Phones. A case study in a Greek University. EAI Endorsed Transactions on Creative Technologies, 5(16). https://doi.org/10.4108/eai.10-4-2018.156382

[3] Xu, G., Yu, Z., Yao, H., Li, F., Meng, Y., & Wu, X. (2019). Chinese text sentiment analysis based on extended sentiment dictionary. IEEE Access, vol. 7, pp. 43749-43762. https://doi.org/10.1109/access.2019.2907772

[4] Shazly, K., Eid, M., & Salem, H., (2020). An efficient hybrid approach for twitter sentiment analysis based on bidirectional recurrent neural networks abstract. International Journal of Computer Applications, 175(17), 32-36. https://doi.org/10.5120/ijca2020920683

[5] A, X. L., A, S. Z., B, P. H., & C, P. L., (2020). Continuous authentication by free-text keystroke based on cnn and rnn. Computers & Security, 96. https://doi.org/10.1016/j.cose.2020.101861

[6] Su, Y., & Kuo, C. C. J. (2019). On extended long short-term memory and dependent bidirectional recurrent neural network. Neurocomputing, vol. 356, pp. 151-161. https://doi.org/10.1016/j.neucom.2019.04.044

[7] Song, T., Sun, J., Zhang, Y., Peng, W., & Song, J. (2020). An RNN Model for Generating Sentences with a Desired Word at a Desired Position. Tehnicki Vjesnik-technical Gazette, vol. 27, no. 1, pp. 81-88. https://doi.org/10.17559/tv-20190929153200

[8] Al-Smadi, M., Talafha, B., Al-Ayyoub, M., & Jararweh, Y. (2018). Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews. International Journal of Machine Learning and Cybernetics. https://doi.org/10.1007/s13042-018-0799-4

[9] Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, Piotr. (2017). Focal Loss for Dense Object Detection. 2017 IEEE International Conference on Computer Vision (ICCV). 2999-3007. https://doi.org/10.1109/iccv.2017.324

[10] A, J. W., A, J. L., A, Z. Y., et al. (2020). Bilstm with multi-polarity orthogonal attention for implicit sentiment analysis. Neurocomputing, vol. 383, pp. 165-173. https://doi.org/10.1016/j.neucom.2019.11.054

[11] Rabovsky, M., & Mcclelland, J. L. (2020). Quasi-compositional mapping from form to meaning: a neural network-based approach to capturing neural responses during human language comprehension. Philosophical Transactions of The Royal Society B Biological ences, vol. 375, no. 1791. https://doi.org/10.1098/rstb.2019.0313

[12] A, H. Q., B, L. W., & C, C. Z., (2020). B&anet: combining bidirectional lstm and self-attention for end-to-end learning of task-oriented dialogue system. Speech Communication, 125, 15-23. https://doi.org/10.1016/j.specom.2020.09.005

[13] Hochreiter, S., & Schmidhuber, J., (1997). Long short-term memory. Neural Computation, vol. 9, no. 8, pp. 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

[14] Palangi, H., Deng, L., Shen, Y., Gao, J., He, X., & Chen, J., et al. (2016). Deep sentence embedding using long short-term memory networks: analysis and application to information retrieval. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 24, no. 4, pp. 694-707. https://doi.org/10.1109/taslp.2016.2520371

[15] Wu, D., & Chi, M., (2017). Long short-term memory with quadratic connections in recursive neural networks for representing compositional semantics. IEEE Access, pp. 1-1. https://doi.org/10.1109/access.2016.2647384

[16] Zhou J, Xu W. (2015) End-to-end learning of semantic role labeling using recurrent neural networks. Proceedings of the Annual Meeting of the Association for Computational Linguistics.

[17] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R., (2012). Improving neural networks by preventing co-adaptation of feature detectors. Computer ence, vol. 3, no. 4, pp. 212-223.

[18] Wei, J. W., & Zou, K., (2019). Eda: easy data augmentation techniques for boosting performance on text classification tasks.

8 Authors

Shen Ji received his B.S. degree in Electronic Information Engineering (School of Information and Safety Engineering) and his M.S. and PH. D. degrees in Spatial Information Science and Technology from Huazhong University of Science and Technology, Wuhan, China. He is currently a lecturer and researcher at Zhongnan University of Economics and Law, Wuhan 430073, China and is actively involved in Deep Learning and Artificial Intelligence.

Tan Fangbi obtained an M.S. degree in Educational Economy and Management from Huazhong University of Science and Technology, Wuhan, China. She is a Ph. D. candidate in Auditing at School of Accounting, Zhongnan University of Economics and Law, Wuhan 430073, China, from 2018. She is currently with Artificial Education, Psychological of College Students, and Sentiment Analysis.