

PAPER • OPEN ACCESS

Chinese Sentiment Analysis of Online Education and Internet Buzzwords Based on BERT

To cite this article: Jibin Zheng *et al* 2020 *J. Phys.: Conf. Ser.* **1631** 012034

View the [article online](#) for updates and enhancements.

You may also like

- [Multi-head self-attention bidirectional gated recurrent unit for end-to-end remaining useful life prediction of mechanical equipment](#)
Changchang Che, Huawei Wang, Xiaomei Ni et al.
- [Lightweight and highly robust memristor-based hybrid neural networks for electroencephalogram signal processing](#)
Peiwen Tong, , Hui Xu et al.
- [Wind turbine blade icing detection using a novel bidirectional gated recurrent unit with temporal pattern attention and improved coot optimization algorithm](#)
Wenhe Chen, Longsheng Cheng, Zhipeng Chang et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Chinese Sentiment Analysis of Online Education and Internet Buzzwords Based on BERT

Jibin Zheng¹, Jinhai Wang^{1*}, Yafeng Ren² and Zeming Yang¹

¹ College of Electronic and Information Engineering, Foshan University, Foshan 528000, China

² Collaborative Innovation Center for Language Research and Service, Guangdong University of Foreign Studies, Guangzhou 510420, Guangdong, China
Email: wangjinhai@whu.edu.cn

Abstract. In recent years, sentiment analysis has made a great breakthrough in the field of natural language processing. Many researchers attempt to build a user comment analysis model of e-commerce products to infer the quality of goods. However, the application of sentiment analysis technology in online education platform and intelligent feedback to students and managers is almost irrelevant. In this paper, we collected a large number of online education platform web resource reviews to build a large corpus to pre-train the Word2vec model for comparative experiments. At the same time, the BERT+BiGRU model is designed for the Feedback System Structure (FSS), which achieves the effect of intelligent interaction between students and teachers and the improvement of the curriculum. On the self-built data set, the BERT+BiGRU model has the best prediction effect, with an accuracy of 98.82%. Compared with traditional machine learning methods such as Naive Bayes, the classification accuracy rate is improved by 21.54%. The experiment result shows that we use BERT+BiGRU on the FSS system to achieve the effect of teacher-student interaction and intelligent feedback of academic intelligence. In addition, we also carry out sentiment analysis on the text with Chinese Internet buzzwords. The experimental results show that the sentiment analysis model of our intelligent feedback system (FSS) has strong analytical capability.

1. Introduction

In China, the most representative ones are MOOC China, China Mo Ke Net, Xuetang Online, etc. According to the 44th Statistical Report on Internet Development in China [1] published by the China Internet Network Information Center (CNNIC). As of June 2019, the number of online education users in China has reached 232 million, an increase of 31.22 million compared with the end of 2018, accounting for 27.2% of the total netizens, of which the number of mobile online education users has reached 199 million, an increase of 530 compared with the end of 2018 Ten thousand, accounting for 23.6% of mobile Internet users. Social media provides a new way to receive news content, users can instantly participate in the spread, interaction and sharing of news events. The scale of online shopping, online finance, and online entertainment users is constantly increasing, and the social platforms provided by these application products have become hot topics and exchange positions. More importantly, people usually express their opinions on the Internet in the form of text and sound, of which text is the most common. At the same time, how to solve the problems such as the low threshold of market access, the unclear demand for online education content, and the need to strengthen network interaction are the problems that need to be considered and solved in the era of intelligent education. The rapid growth of



Internet market and users has brought endless impetus to online education platform. Users' demand for knowledge is constantly deepening, and education platforms with "knowledge less" as products emerge in endlessly. Using emotional analysis to study the feedback data of existing online education users is helpful to find problems and provide new ideas for solving problems. Internet education has changed the original teaching mode and realized the sharing of educational resources, while "Internet Education artificial intelligence" will liberate teachers and online platform managers.

For specific areas, such as online education or social networking, the focus is on the emotional expression of the user. For example, a review of a course: “数据结构是一门非常有用的理论课，可惜这个老师讲得特别糟糕。The data structure is a very useful theoretical course, but this teacher is very bad at it”, which reflect different aspect-based of sentiment analysis (ABSA). As shown in figure 1, “useful and important” is positive emotions, “especially bad”, “influence” are negative emotions. The above analysis shows that, different from ordinary emotional analysis, ABSA relies more on context information and emotional information in different aspects. ABSA can help managers understand users' emotional expression.

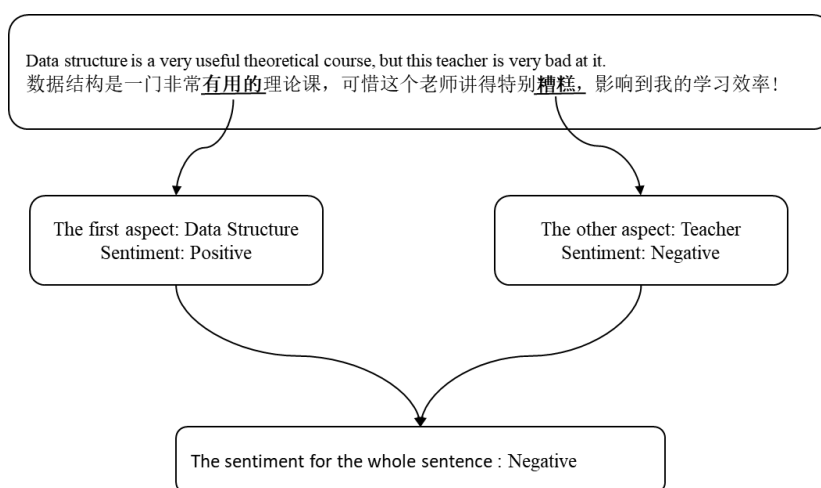


Figure 1. The example of aspect-based of sentiment analysis.

2. Relation Work

In the past decade, the development of emotional analysis has changed from traditional text emotional analysis to modern text emotional analysis. Feng et al. [2] observed that the graphical emoticons are good natural sentiment labels for the corresponding microblog posts and constructed a two-layer graph model using emoji and candidate sentiment words, and selected the top words in the model as sentiment words. Bandhakavi et al. [3] demonstrate how generative unigram mixture model (UMM) based DSEL learnt by harnessing labeled (blogs, news headlines and incident reports) and weakly-labeled (tweets) emotion text can be used to extract effective features for emotion classification. Khoo et al. [4] proposed a general sentiment lexicon called WKWSCI Sentiment Lexicon and compares it with the existing sentiment lexicons. Long et al. [5] used SVM to classify stock forum posts. Azwa Abdul Aziz et al. [6] provide a method known as Contextual Analysis (CA), a mechanism that constructs a relationship between words and sources that is constructed in a tree structure identified as Hierarchical Knowledge Tree (HKT). Machine learning-based methods can automatically extract features, thereby improving classification efficiency. Munezero et al. [7] researched the student learning diary on the online education platform in 2013, tapped the emotional changes in the students' learning process, and provided suggestions for teachers' teaching based on the students' emotional changes to achieve better teaching effect. Xing [8] pointed out the important position of online education in the field of education, but there are emotional barriers separating teachers from students, resulting in the lack of emotions of online learners. He analyzed in detail the reasons for the lack of emotion in online education, and proposed

strategies for achieving emotional interaction. Many researchers have made achievements in the emotional analysis of social platforms or online education platforms. But there are some problems, such as the model structure is very simple, the Chinese Internet buzzwords can not be distinguished, and it is not applied in a complete set of intelligent feedback algorithm or system.

In this paper, we only collect student reviews of computer specialty courses from online education platforms and train our BERT-based sentiment analysis model. Then we apply the trained model to the FSS system architecture we designed. It can analyze students' emotional bias in real time, and cluster these negative emotions or opinions on the course, extract keywords and then sort and push them to the teacher. The teacher gradually improves the course based on the feedback information collected by this system.

3. Method

3.1. Introduction to Data Set

In this paper, the data set is to collect user reviews of computer-related courses of online education platforms and manually label them to train a sentiment analysis model. A total of 25,000 reviews. A total of 25,000 reviews. Partial data sets are shown in table 1. The corpus and data set constructed is a collection of web resources of a large number of online education platforms in Chinese, such as Chinese University MOOC, Tencent Classroom, MOOC, and Xuetang Online.

Table 1. The partial data sets.

Example	Aspect	Sentiment tendency
非常好的课程，比较强调代码的实现过程，对于强化思维非常有帮助。	1. Data structure	Positive
Very good course, more emphasis on the implementation of the code process, for strengthening thinking is very helpful.		
数据结构是一门非常有用的理论课，可惜李老师讲得特别糟糕。	1. Data structure	Negative
Data structure is a very useful theoretical course, but it's a pity that Mr. Li speaks very badly.	2. Mr. Li	

3.2. Word Embedding

Many studies in recent years have pointed out that learning the widely applicable word embedding representation is an inseparable part of modern NLP systems. Pre-trained word embedding is more conducive to subsequent task performance improvement than traditional one-hot encoding import. ELMO [9] generalizes traditional word embedding from different dimensions. By building a two-way neural network with dynamic adjustment of word embedding, it can extract context sensitive features, output pre training word vectors that reflect context semantics, and then solve the problem of polysemy. Compared with ELMO, BERT [10] further broadens the generalization ability of word vector, can fully learn the character level, word level, sentence level and even sentence relationship features, and enhance the semantic representation of word vector, so it shows superior performance than previous methods. As shown in figure 2, the Bert language model is stacked by multiple transformer encoders. Transformer Encoder is mainly composed of Position Encoding, Self-Attention, Layer Normalization + skip connection and activation function. The position information of the initialized word vector X is obtained by linear transformation of sine and cosine functions, where *Loopup* is a dictionary query function used to initialize vectors:

$$X = \text{Lookup}(X) + \text{PositionalEncoding} \quad (1)$$

$$X \in \mathbb{R}^{\text{batch_size} \times \text{seq_len} \times \text{embed_dim}}$$

And then X goes through the attention module:

$$Q = \text{Linear}(X) = X \cdot W_Q$$

$$K = \text{Linear}(X) = X \cdot W_K \quad (2)$$

$$V = \text{Linear}(X) = X \cdot W_V$$

$$X_{\text{attention}} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3)$$

After getting $X_{\text{attention}}$, it reaches the Layer Normalization and skip connection module:

$$X_{\text{attention}} = X + X_{\text{attention}}$$

$$X_{\text{attention}} = \text{LayerNorm}(X_{\text{attention}}) \quad (4)$$

Finally, X_{hidden} is calculated by two-level linear mapping and activation function:

$$X_{\text{hidden}} = \text{Activate}(\text{Linear}(\text{Linear}(X_{\text{attention}}))) \quad (5)$$

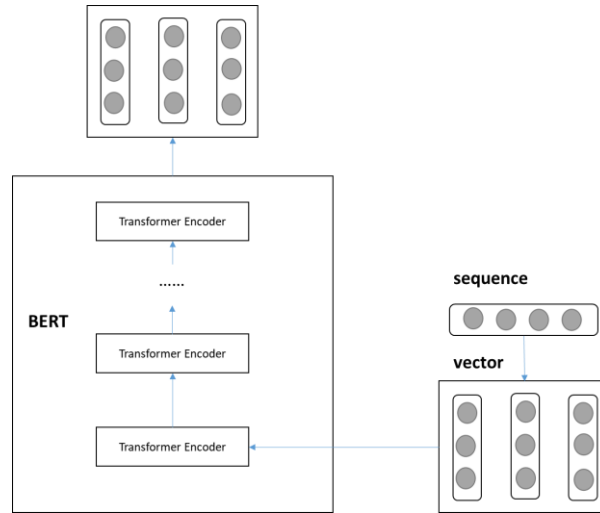


Figure 2. The Bert structure stacked by more transformers.

3.3. The Structure of Sentiment Analysis Model

The sentiment analysis model consists of three components: a sentence encoder and a sentence analyzer. The sentence encoder adopts BERT as the encoder. Then, the BiGRU is used as a downstream neural network for sentiment analysis. This model is applied in the SSUC unit of the FSS system.

4. System Strategy

This part is the core part of the SSUC unit, which is composed of multiple SSUC combined into FSS to realize feedback to the platform. As shown in equation (6), *emotion* is the score of sentiment analysis, t is the length of study time, and *grade* is the final grade.

$$\text{score}_{\text{standard}}^{(k)} = \sigma(\text{emotion})\omega_1 + D(t)\omega_2 + A(\text{grade})\omega_3 \quad (6)$$

As shown in figure 3, the FSS consists of multiple ssuc units with feedback information. The following is the algorithm process of FSS:

(1) Each course corresponds to the SSUC of multiple students, that is, a set of $Course_N$, each element corresponds to a tuple, where neg_seq is the input value of the SSUC.

(2) In the FSS system, all the negative feedback information of each course relation take the student as the primary key, which can get the feedback information set $H_n^{(k)}$ of multiple students, where n is the course number and k is the student number. Each feedback information set $H_n^{(k)}$ is sent to the topic model to extract keywords to obtain the keyword set Ω .

(3) One or more teachers correspond to a student, and after process the set Ω into a string, then write $(teacher, student, keywords)$ to the feedback data table.

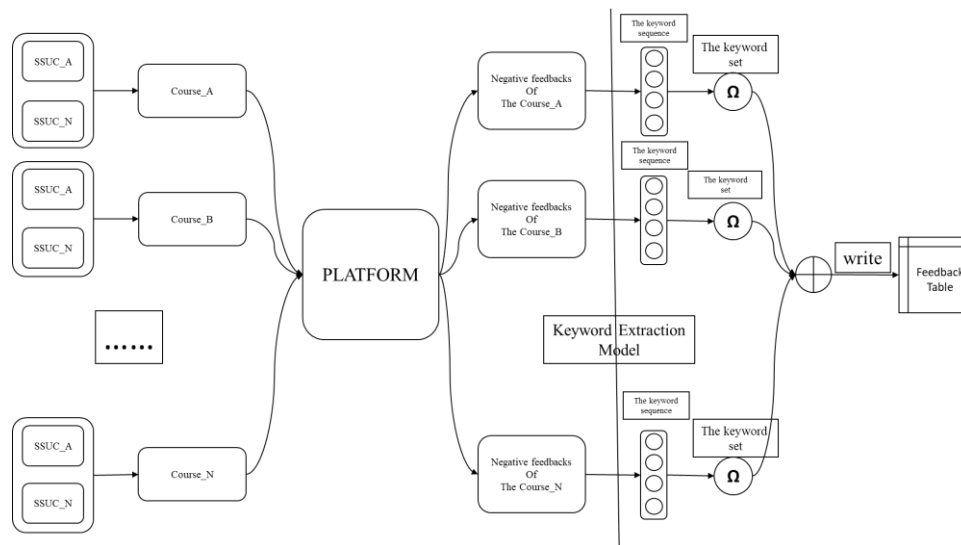


Figure 3. The feedback system structure (FSS).

5. Experiment

5.1. Experimental Environment

In this paper, the experimental environment is shown in table 2.

Table 2. The experimental environment of server

Operating system	Parameters	Operating system	Parameters
CPU	Xeon Silver 4114	Keras	2.2.4
Memory	64G	Python	3.7
GPU	RTX 2080Ti	TensorFlow	1.12.0

5.2. Experimental Parameter Setting and Performance Evaluation

The parameters of Bert Chinese pre training vector are shown in table 3. Other experimental parameters are shown in tables 3 and 4. The ACCURACY is the standard to evaluate the performance of the model in this paper, as shown in equation (7):

$$ACCURACY = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

Table 3. The parameters of BERT model.

BERT-LARGE	Value
layer	24
hidden	1024
heads	16
parameters	330M
Data Source	wiki

Table 4. Model parameter setting.

Parameter	Value
Embedding	300
Bi-GRU of hidden layer nodes	32
Loss	binary_crossentropy
Optimizer	adam
Learning_rate	1e-3
BatchSize	256

5.3. Experimental Result and Analysis

In this paper, efficient and common traditional machine learning methods and deep learning are used to verify the performance, and BERT model and Word2vec language model are compared. Based on the data set constructed, table 5 shows the experimental results of different models. We can see that Bert + bi-gru has the best performance, and the classification accuracy is 98.82%. Compared with word2vec+BI-GRU, the accuracy of Bert model is improved by 4.2%. So bert model has better performance. BERT language model is superior to Word2vec language model when the downstream neural network is completely the same, and the BERT language model is stronger in terms of word meaning expression.

Table 5. Model performance comparison.

Model	ACCURACY	Model	ACCURACY
Word2vec+Naive Bayes	77.28	Word2vec+BI-LSTM	94.40
Word2vec+Logistic Regression	81.48	BERT+ TextCNN	96.21
Word2vec+SVM	85.45	BERT +BI-GRU	98.82
Word2vec+TextCNN	93.81	BERT +BI-LSTM	96.63
Word2vec+BI-GRU	94.61		

5.4. The Sentiment Analysis of Chinese Internet Buzzwords

With the development of Internet in China, the number of Internet users is increasing. Among them, young people use popular words more frequently. For example, “神马东西” is a homonym for “什么东西”, and it means “what thing”. As shown in table 6, in this paper, the sentiment analysis model of our intelligent feedback system has a powerful function of distinguishing Chinese Internet buzzwords.

Table 6. Recognition results of Chinese Internet buzzwords.

Comments	Output probability	Sentiment
这节课内容太顶了，有点好玩 It's a great lesson. It's a bit of fun.	0.99	Positive
这节课讲得太扯了 There are too many mistakes in this lesson.	0.02	Negative
这老师讲课这么无聊？太下饭了 Is the teacher so boring? Teaching ability is too poor.	0.30	Positive

6. Conclusion and Future Work

In this paper, data sets are built and labeled manually, and will be open-source in GitHub later. The experimental results show that FSS can accurately feedback students' learning situation, which benefits from the Bert language model. In addition, our strategy can still accurately predict students' sentiments. Therefore, the model designed in this paper has strong discrimination ability, which can accurately capture the meaning of network words and accurately analyze the sentiment of the final students. We will apply the model to SSUC to realize the function of students' learning feedback. In the future, we will improve the FSS system and try to continue to improve the emotional analysis ability of the model.

Acknowledgment

This work is supported by the National Science Foundation of China (No.61173137), Characteristic innovation projects of Guangdong Universities (2018ktsxcx238), Foshan social science planning project (2019-QN04), Key Project of Foshan Science and Technology Bureau(2018AB003581).

References

- [1] *The 44th Statistical Report on Internet Development in China* <http://www.cnnic.net.cn/hlwfzyj/>.
- [2] Feng S, et al. 2015 A word-emoticon mutual reinforcement ranking model for building sentiment lexicon from massive collection of microblogs *World Wide Web-internet & Web Information Systems* **18** (4) 949-967.
- [3] Bandhakavi A, et al. 2017 Lexicon based feature extraction for emotion text classification *Pattern Recognition Letters* **93** 133-142.
- [4] Khoo C S G and Sathik Basha J 2018 Lexicon-based sentiment analysis: Comparative evaluation of six sentiment lexicons *Journal of Information Science* **44** (4) 491-511.
- [5] Long W, Tang Y R and Tian Y J 2016 Investor sentiment identification based on the universum SVM *Neural Computing and Applications*.
- [6] Aziz A A and Starkey A 2020 Predicting supervise machine learning performances for sentiment analysis using contextual-based approaches *IEEE Access* **8** 17722-17733.
- [7] Munezero M, et al. 2013 Exploiting sentiment analysis to track emotions in students' learning diaries *Proceedings of the 13th Koli Calling International Conference on Computing Education Research*.
- [8] Xing F 2011 Emotion training strategies of online education *2011 International Conference on Electrical and Control Engineering* (IEEE).
- [9] Peters M E 2018 Deep contextualized word representations *arXiv preprint arXiv:1802.05365*.
- [10] Devlin J, et al. 2018 Bert: Pre-training of deep bidirectional transformers for language understanding *arXiv preprint arXiv:1810.04805*.