



Multi-factor evaluation of teaching sentiment analysis in the new era

Yu Zhou^{*a} , Chun Yan^b 

College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao, China

^{*a} zy13255575023@163.com, ^byanchunchun9896@sina.com

Abstract. Teaching reform constitutes a crucial task faced by universities in the new era. This paper analyzes existing issues in online course teaching modes and proposes recommendations for improving these modes. The present study focuses on multifactor evaluation of course instruction, selecting six factors as research subjects from a dataset provided by Kaggle website. We employ sentiment analysis, semantic network analysis, as well as LSTM-based sentiment analysis to delve into implementing online course education from students' perspective while uncovering their concerns and learning needs, ultimately offering relevant suggestions. The conclusions drawn herein possess certain reference value for advancing online education.

Keywords: Teaching evaluation; Affective disposition analysis; Semantic network analysis; LSTM empirical analysis

1 INTRODUCTION

Teaching evaluation is one of the most onerous tasks in various institutions of higher learning, which needs to combine various factors and different standards to provide specific teaching evaluation conclusions for professionals. The existing teaching evaluation is mainly based on an objective scoring mechanism, which lacks effective use of subjective evaluation text information. Most of the evaluation text information is presented in unstructured form. So, sentiment analysis technology is an ideal and effective method to process and analyze subjective evaluation texts [1].

2 LITERATURE REVIEW

Foreign experts have begun to pay attention to and study education and teaching evaluation [2]. Under the influence of foreign teaching theories, China has gradually established a teaching evaluation system. With the continuous advancement of educational reform in our country, the teaching evaluation index system has become more comprehensive. According to the different teaching evaluation methods, the subsequent research has carried out in-depth research on the effectiveness of teaching evaluation

© The Author(s) 2024

R. Magdalena et al. (eds.), *Proceedings of the 2024 9th International Conference on Social Sciences and Economic Development (ICSSSED 2024)*, Advances in Economics, Business and Management Research 289,

https://doi.org/10.2991/978-94-6463-459-4_92

functions and results [3]. At present, with the improvement and development of educational concepts, the construction of a teaching evaluation system pays more attention to the design of evaluation indicators [4]. Scholars at home and abroad have made remarkable achievements in sentiment analysis, and most of them use online review sentiment analysis methods to identify consumers' online review sentiment tendencies [5]. The existing research focuses on the use of semantic analysis technology to conduct emotion mining on the evaluation text, the setting of polarity weight of words, and the calculation of the degree of emotional praise and criticism of the evaluation text [6]. Nowadays, the world tends to be a new era of education, but the research is stagnant on the surface and lacks comprehensive evaluation of the teaching process [7].

It can be seen that there are two main meanings to study the sentiment analysis of curriculum teaching evaluation and its influencing factors in the new era. Firstly, the emotional dictionary of the influencing factors of curriculum teaching evaluation is constructed. Second, it enriches the related research on the influencing factors of teaching evaluation. This study not only enriches the relevant research but also provides new ideas for the relevant research of curriculum teaching. Therefore, this study has important practical significance for promoting the implementation of curriculum teaching in the new era.

3 RESEARCH METHOD

In this study, sentiment analysis, semantic network analysis, sentiment analysis based on the LSTM model, and relevant suggestions are put forward for course teaching evaluation and its influencing factors. Firstly, the relevant literature at home and abroad was reviewed. Second, download the dataset from the Kaggle website and process the data; Then, the corresponding model is established. Finally, the research conclusions are drawn and corresponding countermeasures and suggestions are given.

Considering that Kaggle is a large data set website, this paper has strict requirements on evaluation data, and the retained evaluation data are all real and valuable data with high data quality. Therefore, this paper downloads the student feedback datasets from Kaggle, a public dataset, and analyzes the text data of students' course teaching evaluation on the dataset. The data includes students' evaluation of teaching from six common factors: teaching, course content, extracurricular activities, library facilities, tests, and experiments. This paper visualizes students' educational evaluation data and analyzes the influence of six common factors on teaching evaluation from the perspective of students: teaching, course content, extracurricular activities, library facilities, tests, and experiments. Examples of course teaching evaluation are shown in Table 1.

Table 1. Example of course teaching evaluation

factor	number	comment examples
course content	1	The content of the courses is good but some courses need improvement. Course material is not up to the mark.
	2	The content of courses is average.

	3	All course material provides very good knowledge in depth.
extracurricular activities	1	Complete waste of time. Again, this opinion is strictly personal and may not coincide with others.
	2	No idea about the extracurricular activities.
	3	Extracurricular activities are excellent and provide the best platform for students.
experiment	1	3 hours of practical is so long. It should be 2 hours only.
	2	Sometimes systems at labs are screwed up. Otherwise, everything is at their respective desk.
	3	Practical work is fair.

The content and form of the obtained teaching evaluation data are rich, but they also contain a lot of data with low-value content. If these data are also analyzed, the analysis results will have a large deviation. Then, we should remove the low-quality data, and to ensure the quality of the conclusion data. The data of influencing factors of teaching evaluation of six types of courses were preprocessed respectively, which mainly included text deduplication, cleaning, compression, and deletion. After data preprocessing, a large number of useless data are removed, which provides data quality assurance for subsequent research.

The LSTM model can be used to learn deeper emotional features in teaching evaluation data. The specific calculation formula is shown in Formulas (1) - (8).

$$i_t = \delta(\bar{i}_t) = \delta(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (1)$$

$$f_t = \delta(\bar{f}_t) = \delta(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (2)$$

$$g_t = \tanh(\bar{g}_t) = \tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g) \quad (3)$$

$$o_t = \delta(\bar{o}_t) = \delta(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (4)$$

$$c_t = c_{t-1} \odot f_t + g_t \odot i_t \quad (5)$$

$$m_t = \tanh(c_t) \quad (6)$$

$$h_t = o_t \odot m_t \quad (7)$$

$$y_t = W_{yh}h_t + b_y \quad (8)$$

4 EMPIRICAL ANALYSIS

4.1 Analysis of the emotional tendency of teaching evaluation

This study mainly analyzes the emotional tendency of six common influencing factors. The sentiment score of each review is calculated based on the constructed dictionary, and finally, the review is divided into emotional tendencies according to the sentiment score. The emotional tendency of the review was judged according to its emotional score. If the emotion score is greater than 0, the emotion tendency of the review is

positive, and the review is regarded as favorable. And vice versa. The results of the emotional tendency analysis of the six influencing factors of course teaching evaluation, namely "teaching", "course content", "extracurricular activities", "library facilities", "test" and "experiment", are shown in Table 2.

Table 2. Analysis results of the emotional tendency of course teaching evaluation

factor	positive	neutral	negative	total
teaching	137(74.05%)	35(18.92%)	13(7.03%)	185(100%)
course content	128(69.19%)	27(14.59%)	30(16.22%)	185(100%)
extracurricular activities	154(83.24%)	19(10.27%)	12(6.49%)	185(100%)
library facilities	130(70.27%)	24(12.97%)	31(16.76%)	185(100%)
test	131(70.81%)	30(16.22%)	24(12.97%)	185(100%)
experiment	132(71.35%)	16(8.65%)	37(20%)	185(100%)

The results show that students' satisfaction with all six factors is high. Among them, 74.05% of students rated "teaching" positively, 18.92% rated moderately, and 7.03% rated poorly. The positive rating for "course content" was 69.19%, with a moderate rating of 14.59% and a poor rating of 16.22%. The positive rating for "after-school activities" was 83.24%, with a moderate rating of 10.27% and a poor rating of 6.49%. The positive rating for "library facilities" was 70.27%, with a moderate rating of 12.97% and a poor rating of 16.76%. The positive rating of "quizzes" was 70.81%, the moderate rating was 16.22%, and the poor rating was 12.97%. The positive rating of "experiments" was 71.35%, the moderate rating was 8.65%, and the poor rating was 20%.

If there are both positive comments and neutral or negative comments in a review, the positive comments and neutral or negative comments in the same review are separated and grouped into positive comments and neutral or negative comments respectively for analysis, to achieve a more detailed division of the review and more comprehensive and accurate mining of its relevant information. "Pattern and procedure of the examination is very good. Paper checking needed to be improved." Divided into "Pattern and procedure of the examination is very good." and "Paper checking needed to be improved." It also classified "Pattern and procedure of the examination is very good." as positive, and "Paper checking needed to be improved." as neutral or negative.

Firstly, the word frequency statistics of the positive and neutral/negative evaluation of the six types of factors affecting the course teaching evaluation were carried out respectively, and then the word cloud maps of the positive and neutral/negative evaluation of the six types of factors affecting the course teaching evaluation were drawn based on the word frequency statistics. In the word cloud map, words with more occurrences have larger shapes. The drawing of word cloud map can not only intuitively show the effect of word segmentation, but also show students' concerns.

For example, in the evaluation word cloud map of factor "test," it can be seen that high-frequency words include good, pattern, checking, examination, excellent, exam, marks, paper, distribution and so on. A similar analysis was made in other factor analyses. The results of emotional orientation analysis showed that the praise rate of the six

factors reached more than 70%. The word cloud map shows that the text word segmentation has achieved good results, and intuitively shows the positive and neutral/negative high-frequency words of the six categories of factors, which reflects the students' concerns to a certain extent.

4.2 Semantic network analysis

In recent years, semantic network has become one of the most widely used methods in artificial intelligence such as natural language processing, which is of great value to the acquisition of semantic information.

The results show that students think the content of the test is substantial and comprehensive, to test their mastery of the theoretical content of the course and deepen their memory of the theoretical content with the help of the test questions. The teacher sets the test questions with moderate difficulty and controls the time for students to answer. The main problem with the test is that the style of the test questions is too single, most of them are objective choice questions, and there are few open questions. Although it can reduce the task of teachers' assessment to a certain extent, it is not conducive to examining students' comprehensive induction and analysis ability. The difficulty of the test questions set by the teacher is appropriate, which can examine the degree of student's mastery of knowledge points to a certain extent, but the degree of differentiation of the test questions is not large for students who have fully mastered the knowledge points. Other factor analysis is given similarly.

4.3 Teaching evaluation analysis based on the LSTM model

To a certain extent, the analysis verifies the robustness of the data and the universality of the analysis conclusions, and provides scientific and reasonable countermeasures and suggestions.

Table 3. Generated dictionary

Num.	word	Num.	word	Num.	word	Num.	word
1	good	6	in	11	and	16	us
2	the	7	work	12	of	17	evaluations
3	lab	8	practical	13	very	18	that
4	is	9	not	14	it	19	they
5	are	10	to	15	labs	20	knowledge

Firstly, with the help of Python software, the toolkit was imported to read the training and test data, and the factor "experimental" data totaled 185 pieces. Secondly, the teaching evaluation text is pre-processed and standardized. Thirdly, the words in the text are mapped to numbers, and the token in keras is used to achieve this. Finally, the pad_sequences method is used to complete the same length. The dictionary (top 20 items) is obtained by numbering the words from highest to lowest frequency in Table 3.

Through the previous data preprocessing, the content of each comment is turned into 100 numbers that can be read by the computer. The purpose is to realize the embedding

of words into a multi-dimensional matrix, so that the words with similar semantics are closer in spatial distance. The input dimension is 400 and the length of each review data is 100. The dropout probability of 0.2 is specified in the dropout layer, and 20% of the neurons in the previous layer are randomly discarded to make the model robust to some degree to prevent over-fitting of the sample data.

Adam optimizer is used to train the algorithm, and the weights of each connection layer are optimized. The parameters are set as follows: learning efficiency is 0.0001, infinity is 10-8, and the attenuation factor of learning rate in each round of training is 0.0. A decay factor of 0 means that the learning rate remains constant during training. During the experiment, the labeled corpus was divided into training set and validation set according to 8:2.

The training set and validation set in the above training process are visualized respectively, as shown in Figure 1.

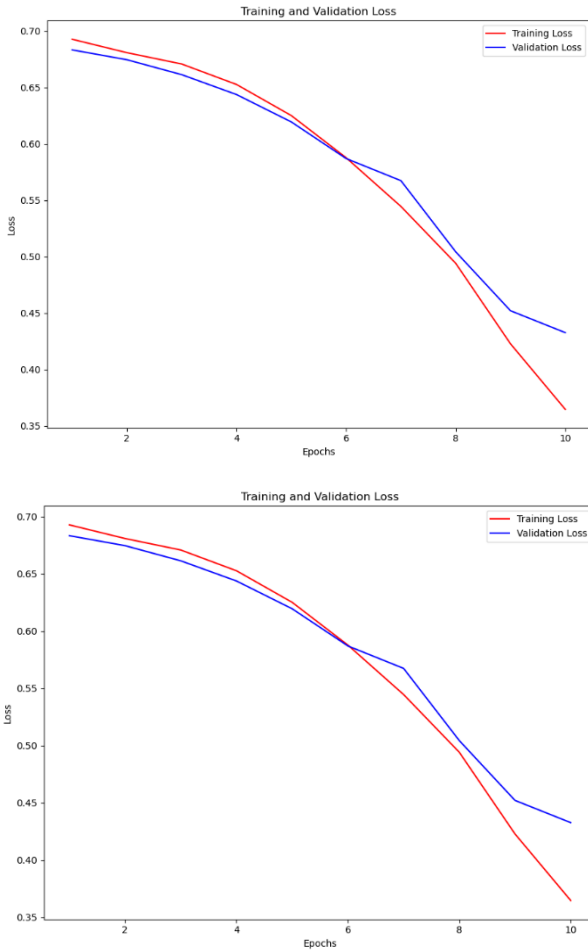


Fig. 1. (a) training and validation loss (b) training and validation accuracy

The visualization results show that the number of iterations is set to 10, and the loss value of the training set during each iteration process decreases as the accuracy decreases. In the validation set, the loss function first decreases with the decrease of the accuracy rate, then increases with the increase of the accuracy rate, and finally tends to be stable. At the seventh iteration, the training and validation results are the best. The training set exhibits a loss value of 0.5635, accompanied by an accuracy rate of 0.7209. Similarly, the validation set demonstrates a loss value of 0.5243 and an accuracy rate of 0.6607.

5 CONCLUSIONS AND SUGGESTIONS

In this study, six factors affecting educational evaluation on the data set of Kaggle website are selected as the research objects, and the analysis shows that (a) the course teaching evaluation is recognized and affirmed by students. (b) Students mainly focus on course content, teaching mode and learning duration. (c) Combining advanced technology with innovative teaching models, such as video captioning. The solution of curriculum problems and the realization of learning need the support of related technologies.

REFERENCES

1. Wang, B.H., et al. (2021) Sentiment Analysis of Student Teaching Evaluation Based on Deep Learning. *e-Education Research (Chinese)*, 42(04): 101-107. DOI: 10.13811/j.cnki.eer.2021.04.014.
2. Dianne M. Timm. (2008) Developing Outcomes-Based Assessment for Learner-Centered Education: A Faculty Introduction (review). *Journal of College Student Development*, 49(4): 392-394. <https://doi.org/10.1353/csd.0.0023>.
3. Luo Y.P., et al. (2018) Design and Application of Teaching Evaluation System Based on Sentiment Mining. *China Educational Technology*, (04): 91-95. <https://link.cnki.net/urlid/11.3792.G4.20180412.0958.026>.
4. Wu L.B., et al. (2021) Reform and Practical Paths of Classroom Teaching Evaluation Under Artificial Intelligence. *China Educational Technology*, (05): 94-101.
5. Basiri E. M., Nemati S., Abdar M., et al. (2021) ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. *Future Generation Computer Systems*, 115: 279-294. <https://doi.org/10.1016/j.future.2020.08.005>.
6. Pong-inwong C., Songpan W. (2019) Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining. *International Journal of Machine Learning and Cybernetics*, 10(8): 2177-2186. <https://doi.org/10.1007/s13042-018-0800-2>.
7. Onan, A. (2020) Mining opinions from instructor evaluation reviews: A deep learning approach. *Computer Applications In Engineering Education*, 28(01): 117-138. Doi: 10.1002/cae.22179.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

