An Automatic Analysis and Evaluation System Used for Teaching Quality in MOOC Environment

Sicheng Yang
School of Automation
Beijing Institute of Technology
Beijing, P. R. China
yangsicheng81@bit.edu.cn

Yaping Dai
School of Automation
Beijing Institute of Technology
Beijing, P. R. China
daiyaping@bit.edu.cn

Simin Li
School of Automation
Beijing Institute of Technology
Beijing, P. R. China
3120190899@bit.edu.cn

Kaixin Zhao
School of Automation
Beijing Institute of Technology
Beijing, P. R. China
zhaokaixini@bit.edu.cn

***Abstract*—To solve the problem of automatically analyzing and evaluating the teaching content and effect of teachers in massive open online courses (MOOC) environment, an automatic teaching evaluation system is proposed in this paper to evaluate the sentiment of teacher and content of "online classes". Firstly, the multimodal sentiment analysis model based on voice and text is built, which can determine the degree of "positive" and "negative" sentiments of teachers. Then, the textbook and Baidu Encyclopedia are used as two kinds of syllabus. The "3D matching degree decision model" is built to compare the differences between the teaching content and the syllabus, then the matching degree of teaching content is given. According to the results of the sentiment analysis and matching with syllabus, the teaching quality can be effectively judged. Finally, experiments in are conducted in MOOC environment. The results of the automatic analysis and evaluation system used for teaching quality perform well.**

***Index Terms*—teaching content modeling, massive open online courses (MOOC), sentiment analysis, text match**

I. Introduction

Massive open online courses (MOOC) [1], with the freedom of time and low cost, has become the current trend [2]. But there are very few researches on the automatic analysis of the effectiveness and teaching quality in MOOC environment. At present, the evaluation of course effectiveness in MOOC environment almost comes from the manual evaluation of experts and students. That is more subjective and less efficient. Moreover, in online education, there is a spatial separation between teachers, students and listening experts, which leads to the students and listening experts cannot provide feedback to teachers timely in each lesson. And students often cannot complete their evaluations until the course has been fully completed, which is not conducive to teacher to reflect on their teaching effectiveness timely after each lesson. Therefore, there is an urgent need for an objective, efficient and timely evaluation for teaching quality in MOOC environment. So that after each lesson, teachers are more likely to "meet the requirements of the syllabus" and "express positive emotions", which also improves the quality of course and attracts the attention of students in MOOC environment.

This work was supported by the Beijing Municipal Natural Science Foundation under grant No. L191020 and No. 3192028, the National Talents Foundation of China under grant No. WQ20141100198.

In this paper, by means of artificial intelligence theoretical methods and techniques, an automatic analysis and evaluation system is proposed, which is able to analyze the teaching content and provide feedback to the teacher timely on whether the teaching content meets the syllabus requirements and whether the sentiment of teacher is positive. Specifically, the video of MOOC course is converted into audio and text. The sentiments of the audio and text are analyzed separately. Then the fusion algorithm is used to make decision. The model for the sentiment analysis of the teacher is built. Even more, the video and the syllabus are matched by the "3D matching degree decision model". Then the final matching degree of the teaching content and the syllabus is obtained by data fusion. In the end, the system produces the 'Teaching Evaluation' report automatically with confidence coefficient of positive sentiment and the matching degree of teaching content and syllabus. The overall block diagram of the system is shown in Fig. 1.

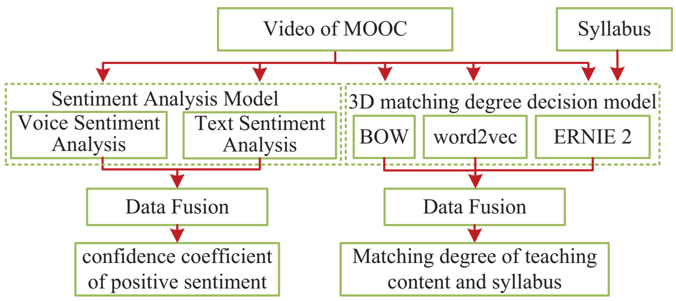


Fig. 1. Overall block diagram.

The paper is organized as follows. In Section II, the audio and text from MOOC video are extracted, and the multimodal sentiment analysis model for teaching content is buildt to give the Confidence Coefficient of Positive Sentiment (CCPS) by multimodal fusion. In Section III, the syllabus is constructed, and the "3D matching degree decision model" is built to give the Matching Degree of the Teaching content and the Syllabus (MDTS) by inter-model and inter-syllabus fusion. In Section IV, the above algorithm is demonstrated in MOOC environment. Section V gives the conclusion and outlook. The full dataset, codes and results are available for use at https://github.com/YoungSeng/DTPI2021.

II. The Model of Multimodal Sentiment Analysis for Teaching Content

The proposed model of multimodal sentiment analysis for teaching content in MOOC environment is shown in Fig. 2.

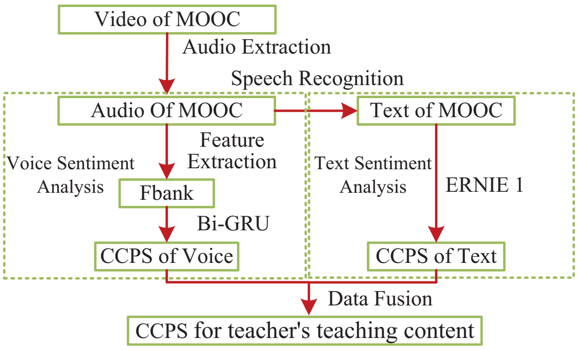


Fig. 2. The multimodal sentiment analysis model.

A. Voice Sentiment Analysis Model

1) *Feature Extraction:* Log Mel filterbank coefficients (Fbank [3]) is used as the feature of speech signals. After calculation [4], the Fbank features are obtained.
2) *Bidirectional Gate Recurrent Unit(Bi-GRU [5]):* The hidden layer outputs through the fully connected layer. The positive sentiment result of softmax binary classification is used as CCPS of voice noted as P_1.

B. Text Sentiment Analysis Model

This paper focuses on Chinese MOOC environment. So the model for text sentiment analysis is based on Enhanced Representation from kNowledge IntEgration (ERNIE [6]) proposed by Baidu. After rounds of fine-tuning, the positive sentiment result of ERNIE 1 softmax binary classification is used as CCPS of text noted as P_2.

C. Sentiment Analysis based on Bayesian reasoning

Based on Bayes' rule, the probability of event H_i occurring given evidence E can be expressed as:

$$P(H_i|E) = \frac{P(E|H_i)P(H_i)}{\sum_i [P(E \mid H_i) P(H_i)]}, \tag{1}$$

where $P(H_i|E)$ denotes the posterior probability of the occurrence of event H_i given evidence E, $P(E|H_i)$ denotes the probability of the occurrence of evidence E given the occurrence of event H_i, $P(H_i)$ denotes the prior probability of the occurrence of event H_i, and $\sum_i [P(E \mid H_i) P(H_i)]$ denotes the full probability of evidence E given the occurrence of each event.

Assume that the probabilities of teaching content with positive and negative sentiments are $P(O_1)$ and $P(O_2)$, respectively. Assume that the 'POS' is the positive sentiment judged by the model, and CCPS is the probability that the model is 'POS' under the condition that the sentiment of the teaching content is positive. Then the probability of the voice is 'POS' under the condition that voice is positive is $P_{\text{voice}}(\text{POS}|O_1) = P_1$ and the probability of the text is 'POS' under the condition that the text is positive is $P_{\text{text}}(\text{POS}|O_1) = P_2$. Assum that the voice and the text are independent of each other, the formula of joint probability of the voice and text are 'POS' are expressed as follows:

$$\begin{aligned} P(\text{POS},\text{POS}|O_1) &= P_{\text{voice}}(\text{POS}|O_1)P_{\text{text}}(\text{POS}|O_1) \\ &= P_1P_2, \end{aligned} \tag{2}$$

$$\begin{aligned} P(\text{POS},\text{POS}|O_2) &= P_{\text{voice}}(\text{POS}|O_2)P_{\text{text}}(\text{POS}|O_2) \\ &= (1-P_1)(1-P_2). \end{aligned} \tag{3}$$

Substitute (2) and (3) into (1), the final CCPS of teaching content can be calculated by the following equation:

$$P(O_1|\text{POS},\text{POS}) = \frac{P_1P_2P(O_1)}{P_1P_2P(O_1) + (1-P_1)(1-P_2)P(O_2)}. \tag{4}$$

III. The Matching Model of Teaching Content and Syllabus

For each sentence of the text of teaching content, the matching degree between it and each sentence of the syllabus is calculated. The highest result is selected as the matching degree of that sentence and the syllabus. The matching model of teaching content and syllabus in MOOC environment is shown in Fig. 3.

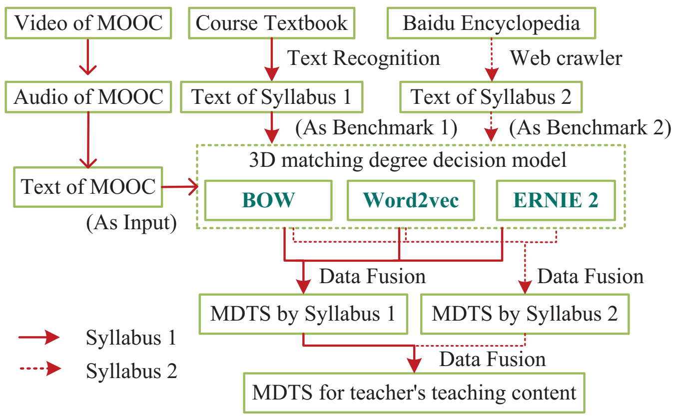


Fig. 3. The matching model of teaching content and syllabus.

A. Definition of Two Types of "Syllabi"

Syllabus 1 is the textbook corresponding to the course in MOOC Syllabus 2 is searched from Baidu Encyclopedia using the course name as the keyword.(optional, can also be replaced with other syllabus as benchmark)

B. Build the "3D Matching Degree Decision Model"

1) *Bag of Words (BOW [7]) Model:* This dimension ignores the word order, syntax and grammar of the text. The cosine similarity of the two vectors is used as MDTS calculated by this dimensional model noted as P_a.
2) *Word2vec [8] Model:* This dimension focuses on the relationships between words. The cosine similarity of

the two vectors is used as MDTS calculated by this dimensional model noted as P_h.

3) *ERNIE 2 Semantic Matching Model:* This dimension focuses on the relationships between sentences. After rounds of fine-tuning, the matched result of ERNIE 2 softmax binary classification is used as MDTS calculated by this dimensional model noted as P_c.

C. Matching Degree Analysis Based on Bayesian Reasoning

Assume that the probabilities of teaching content matching and not matching the syllabus are $P(M_1)$ and $P(M_2)$ respectively. Assume that the 'Match' is the case where the 1D model is judged to be match, and MDTS calculated by each dimension model is the probability that the model is 'Match' under the condition that the teaching content matches with the syllabus. Then, the probability of teaching content is 'Match' for each dimension model are given as follows: $P_{\text{BOW}}(\text{Match}|M_1) = P_a$, $P_{\text{Word2vec}}(\text{Match}|M_1) = P_b$, and $P_{\text{ERNIE}}(\text{Match}|M_1) = P_c$. Assume that the three models are independent of each other, the joint probability of the teaching content and syllabus are 'Match' in three dimensions are expressed as follows:

$$P(\text{Match}, \text{Match}, \text{Match}|M_1) = P_a P_b P_c, \tag{5}$$

$$P(\text{Match}, \text{Match}, \text{Match}|M_2) = (1-P_a)(1-P_b)(1-P_c). \tag{6}$$

Substitute (5) and (6) into (1), the MDTS of teaching content using the "3D matching degree decision model" after data fusion can be calculated by the following equation:

$$P(M_1|\text{Match}, \text{Match}, \text{Match}) = \frac{P_a P_b P_c P(M_1)}{P_a P_b P_c P(M_1) + (1-P_a)(1-P_b)(1-P_c)P(M_2)}. \tag{7}$$

The MDTS by syllabus 1 and 2 can be calculated using (7), and the results noted as P_A and P_B, respectively. Assume that the 'Match3' is the case where the 3D model is judged to be match. Substitute P_A and P_B into (1), the final MDTS for teaching content is expressed as follows:

$$P(M_1|\text{Match3}, \text{Match3}) = \frac{P_A P_B P(M_1)}{P_A P_B P(M_1) + (1-P_A)(1-P_B)P(M_2)}. \tag{8}$$

IV. Experimental Results and Analysis

In the experiment, the computer processor is Intel(R) Core(TM) i5-7300HQ and 8 GB RAM. The GPU is RTX 2080Ti. The software environment is Anaconda 4.10.0, and PyCharm Community Edition 2020.2.3 with Python 3.8.5.

A. Experiment Preparation

1) *Video of MOOC:* The audio of MOOC is extracted through the format factory. The speech transfer API of iFlytek open platform is called to translate the audio into the corresponding text.

2) *Syllabus:* Adobe Acrobat Pro DC is used to convert the textbook in PDF format into picture format. Baidu AI open platform scene text recognition API is called to convert the picture into corresponding text. Crawler is used to get the text of Baidu Encyclopedia.

The prior probabilities are set to $P(O_1) = 0.8$, $P(O_2) = 0.2$, $P(M_1) = 0.5$, and $P(M_2) = 0.5$, respectively.

B. Example 1 of Experiment in MOOC Environment

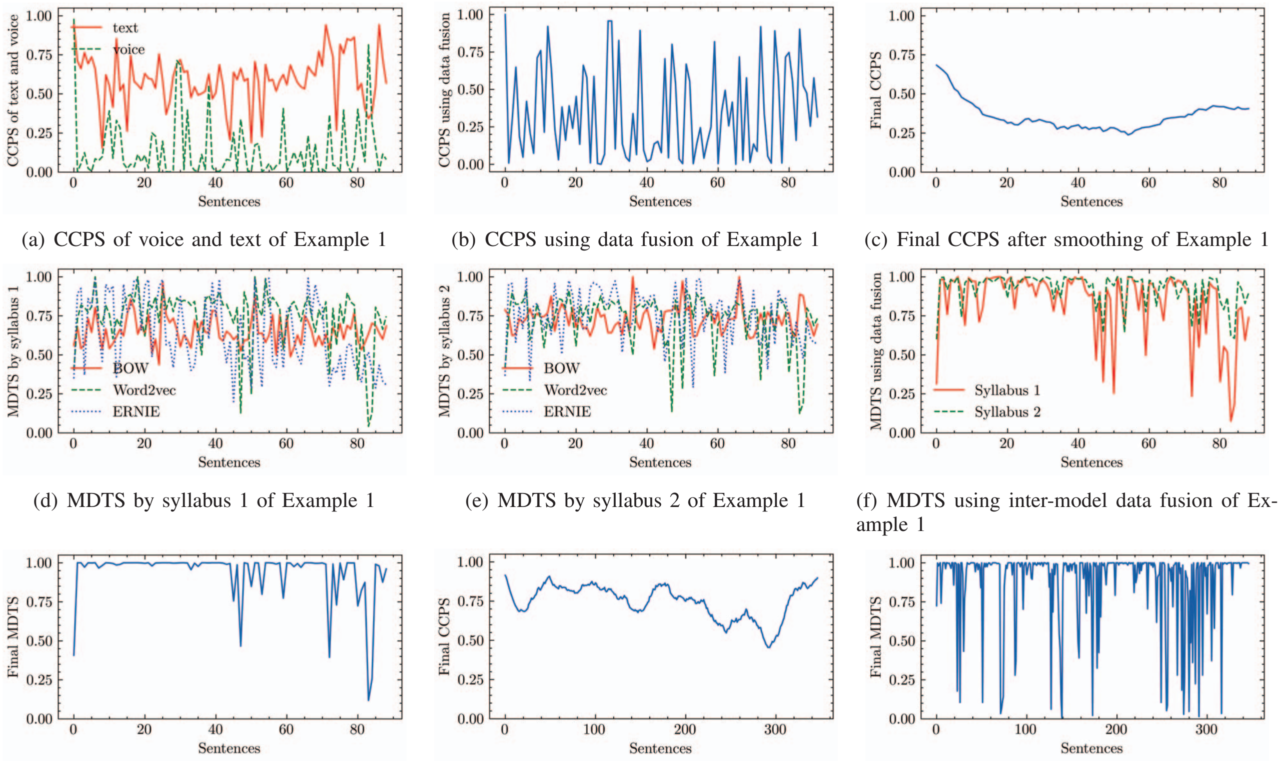
The experiment is conducted with the course "Analysis of Typical Problems in College Physics-Mechanics and Thermodynamics": "A Brief Review of Newton's Theorem of Motion" on the website"Chinese University MOOC" as Example 1. Syllabus 1 is from pages 25-29 of "University Physics (Volume 1) Mechanics and Thermodynamics". Syllabus 2 is crawled through Baidu Encyclopedia with the keyword "Newton's Laws of Motion". There are 89 sentences in this course. The details of the first two sentences, "Hello, everyone" and "This week we are doing exercises related to Newton's laws of motion" are shown in Table I.

TABLE I
DETAILED RESULTS FOR THE FIRST TWO SENTENCES OF EXAMPLE 1

Evaluation	Model Items		Sentence 1 Results	Sentence 2 Results
CCPS	Text		0.9493	0.7098
	Voice		0.9806	0.0009
	Final Fusion results		0.9997	0.0084
MDTS	***Matching Degree with Syllabus 1***	BOW	0.5603	0.6787
		Word2vec	0.4023	0.7292
		ERNIE 2	0.3481	0.8926
		Fusion	0.3141	0.9793
	Matching Degree with Syllabus 2	BOW	0.7883	0.7532
		Word2vec	0.4121	0.7374
		ERNIE 2	0.3648	0.9522
		Fusion	0.5999	0.9942
	Final Fusion results		0.4071	0.9999

On the time domain of a lesson, the CCPS of the text and voice are shown in Fig. 4(a). Using (4) for data fusion between text and voice, the result is shown in Fig. 4(b). In order to visualize the sentiment change of the teaching content, assuming that the sentiment of the teaching content is continuously changing. Savgol filter is used for smoothing with the window length set to 51 and the order of the polynomial fit set to 3. The result curve is shown in Fig. 4(c). The MDTS by syllabus 1 and syllabus 2 calculated by the "3D matching degree decision model" are shown in Fig. 4(d), and Fig. 4(e), respectively. MDTS by the syllabus 1 and syllabus 2 using (7) for inter-model fusion are shown in Fig. 4(f). Final MDTS using (8) for inter-syllabus fusion is shown in Fig. 4(g).

Teaching Evaluation: The CCPS for teaching content fluctuated around 0.4, which can be considered as negative sentiment. The teacher greeted to student in class and arranged after-class tasks at the end of class, so the MDTS decreased a little The overall teaching content matched the syllabus very well. Further analysis, the CCPS of teacher's voice was low, which may be related to the teacher's own timbre and tone.



(a) CCPS of voice and text of Example 1 (b) CCPS using data fusion of Example 1 (c) Final CCPS after smoothing of Example 1

(d) MDTS by syllabus 1 of Example 1 (e) MDTS by syllabus 2 of Example 1 (f) MDTS using inter-model data fusion of Example 1

(g) Final MDTS of Example 1 (h) Final CCPS of Example 2 (i) Final MDTS of Example 2

Fig. 4. Experimental result curves in MOOC environment.

C. Example 2 of Experiment in MOOC Environment

The experiment is conducted with the course “Introduction to Chinese Culture”: “1.1 What is “Culture”” on the website“Chinese University MOOC” as Example 2. Syllabus 1 is from pages 1-3 of “Introduction to Chinese Culture, 2nd edition”. Syllabus 2 is crawled through Baidu Encyclopedia with the keyword “Culture”. There are 347 sentences in this course, the CCPS after smoothing (parameters of Savgol filter are same as Example 1) is shown in Fig. 4(h). The Final MDTS is shown in Fig. 4(i).

Teaching Evaluation: The CCPS for teaching content fluctuated around 0.8, so the teaching content can be considered positive. The MDTS fluctuated widely, and further analysis showed the following reasons: first, the teacher’s pronunciation was not standard; second, the teacher had a lot of pauses while teaching; third, the teacher taught many examples and stories while teaching.

V. Conclusion

To solve the problem of automatically analyzing and evaluating the teaching content and the effects of the course in MOOC environment, an system to evaluate sentiment and content of courses automatically is proposed in this paper. The system analyze voice and text based on the videos of MOOC and syllabus,and produce ”Teaching Evaluation” report automatically with CCPS and MDTS.

We find there are some aspects worthy of improvement in the system, such as (1) The categories of sentiment are sketchy. The sentiment contained in a voice or a text may not be simply classified as positive or negative. (2) Since the matching is based on sentence pairs, the whole relationship between the teaching content and the syllabus may be ignored. Thus, in order to develop a more accurate system for analyzing and evaluating the teaching quality, fine-grained sentiments can be analyzed and the way of matching can be improved, which is also the direction of the follow-up work of this study.

References

[1] J. Haber, MOOCs. Cambridge, Mass.: MIT Press, 2014, pp. 37–39.

[2] Y. DAI, F. YANG, H. ZHAO Z. JIA and K. HIROTA, “Auto Analysis System of Students Behavior in MOOC Teaching,”, Acta Automatica Sinica, vol. 46, no. 4, pp. 681-694, 2019.

[3] M. Bahoura, ”Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes”, Computers in Biology and Medicine, vol. 39, no. 9, pp. 824-843, 2009.

[4] A. Mohamed, G. Hinton and G. Penn, ”Understanding how deep belief networks perform acoustic modelling”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012, pp. 4273-4276.

[5] K. Cho et al., ”Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”, Computer Science, 2014.

[6] Y. Sun et al. ”ERNIE: Enhanced Representation through Knowledge Integration”, CoRR, vol. abs/ 1904.09223, Apr 2019.

[7] Y. Zhang, R. Jin and Z. Zhou, ”Understanding bag-of-words model: a statistical framework”, International Journal of Machine Learning and Cybernetics, vol. 1, no. 1-4, pp. 43-52, 2010.

[8] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” CoRR, vol. abs/1301.3781, May 2013.