Research on learner feedback and text analysis of web design and development course—Based on Online Open Course Reviews from 6 Institutions

Yi Meilian

Guangzhou College of Applied Science and Technology, Guangdong, Guangzhou, 511370, China

Abstract: In order to understand the learners' needs, concerns, and directions for improving web design and development courses, this study collected similar courses from six universities in China's MOOC (Massive Open Online Course) platform(https://www.icourse163.org/). A combination of quantitative and qualitative analysis methods was employed, utilizing Python for word frequency analysis and ROSTCM6 for sentiment analysis. In-depth analysis was conducted on highly praised and negative comments. The experimental results revealed that the evaluation of the courses mainly focused on course content and instructor performance, with significantly more positive feedback than negative feedback. Students expressed concerns regarding the lack of practical materials, variations in software versions, and excessive theoretical content during lectures. Based on the research findings, recommendations were proposed in five aspects, including course content organization and arrangement.

Keywords: MOOC; Course review; Text analysis

1. Introduction

The term Massive Open Online Course (MOOC) emerged in 2008 and became an educational buzzword in 2012, receiving more and more attention [1].MOOC platforms provide large-scale open online courses for free via the Internet, which not only provide high-quality courses and learning opportunities for learners worldwide, but also MOOC platforms provide opportunities to improve educational equity and traditional teaching. The rise of MOOCs has promoted the development of online education and has had an important impact on traditional higher education, but there are also quality challenges such as low completion rates and difficult evaluation. Analysis of online course reviews can provide insight into learner satisfaction, learner concerns, and focus analysis through negative evaluations can provide assistance for future course improvement.

2. Related Studies

Course reviews are the comments and feedback provided by learners on courses offered on online open platforms, documenting their learning experiences and evaluations. The information contained in these reviews is of great value for understanding learner needs, improving course quality, and optimizing the learning experience. Therefore, review mining has always been a hot topic in MOOC-related research.

Course review mining is mainly based on technologies such as text mining, natural language processing, and social network analysis. In text mining and language processing, researchers collect textual reviews, preprocess them, extract features from the review texts, perform sentiment analysis, and mine topic-oriented reviews [2,3]. The most commonly used method for topic mining is Latent Dirichlet Analysis (LDA). Through text mining and sentiment analysis, learner satisfaction with the course can be understood [4]. Social network analysis involves metrics such as nodes, edges, clustering coefficients, and centrality, which can be used to analyze the relationships and interaction network structure among learners. It helps identify key actors and core members in the network, detect interaction patterns among learners, evaluate the effectiveness of learning communities, and assess the learning atmosphere [5].

Related research on mining learner reviews indicates a significant correlation between learner feedback and course evaluation. Positive reviews are positively correlated with learning outcomes, and the sentiment expressed in learner comments can predict their satisfaction and effectiveness [6]. The satisfaction of MOOC learners is primarily influenced by the learning experience, and an ideal learning experience can lead to higher learner satisfaction ratings [7]. Generally, the stronger the interactivity in

the comments, the higher the course evaluation. Interactive comments can also predict the attractiveness of the learning content.

Although there are numerous existing literature on the mining of course reviews, most of the current research focuses on analyzing and studying individual courses, with relatively few comparative analysis studies. This study aims to conduct a comparative analysis by examining courses of the same type from different schools. By mining course reviews, we intend to discover learner evaluations and needs, delve deeper into learner opinions, and provide support for future course improvements and instructional decision-making.

3. Research Design

3.1. Research Logical Framework

China University MOOC offers courses from major institutions to the general public is the largest online course platform in China with the largest number of courses and the largest number of registrants, and it is a typical representative of online open courses. Therefore, this study uses the learner review data of Chinese university MOOC, mainly through the analysis and mining of Chinese university MOOC course reviews to discover the content of learners' emotional tendency, demand and evaluation information of the course, and the research framework is shown in Figure 1.
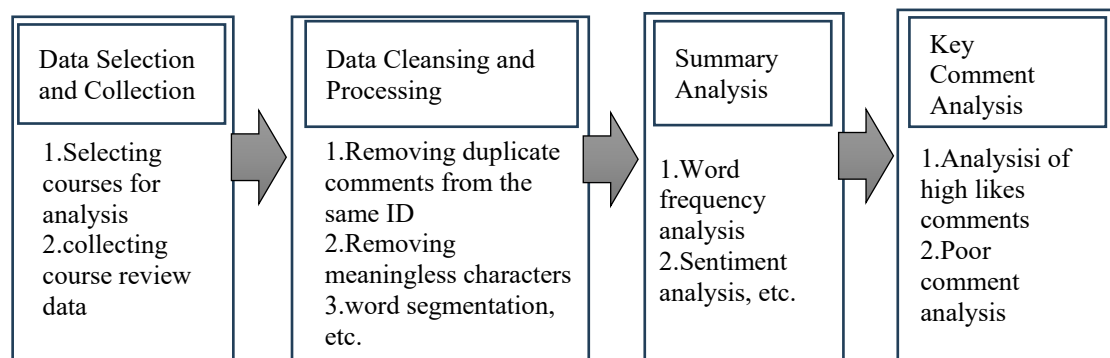

Figure 1. Thesis research framework

3.2. Research Methods

In Chinese universities' MOOC platforms, courses related to "web design," "web development," and "website creation" were searched using these keywords. Courses with a review count of over 200 were selected. The data collection was done using a tool called "Octoparse," which provides easy and efficient access to course review data on the Chinese universities' MOOC platform. The collected data was then processed by removing duplicates, eliminating meaningless comments with repetitive IDs or special characters (such as "123456" or "!!!!!!"), to prepare for the subsequent analysis. For word segmentation and frequency analysis, the Python library "jieba" was used, along with a custom dictionary specific to this segmentation task. The "Harbin Institute of Technology" stop word list was used with slight modifications. During the frequency analysis, only nouns, proper nouns, and gerunds were retained, and words with a length of two or more characters were considered. Sentiment analysis was performed using the ROSTCM6 text mining tool. In the detailed analysis, particular emphasis was placed on extracting highly praised comments and negative evaluations for further examination.

4. Empirical Analysis

4.1. Data Description

This study uses Octopus data collection tool to collect the review data of Chinese university MOOC web design and production courses, which are highly practical courses and are permanent courses for computer, software engineering, e-commerce, design majors, and are involved in both arts and science disciplines. At present, more than twenty institutions have offered related courses on this platform, and this study selects courses with more than 200 reviews, collected on May 12, 2023, and the collected data and the cleaned data information are shown in Table 1.

Table 1. Basic information table of data collection

Course Serial Number	Course Name	School	Number of participants in the course	Number of Course Reviews (Articles)	Number of Valid Course Reviews (Article)
1	Web Design and Production	Xianyang Vocational and Technical College	1721	366	337
2	Web Design Fundamentals	Guangdong Polytechnic of Industry and Commerce	1744	364	324
3	Web Design Fundamentals	Beijing Language and Culture University	3118	282	277
4	Web Site Development	Wuhan Technical College of Communications	620	275	255
5	Web Design Client Technology	Changsha Social Work College	545	257	247
6	Web Design and Development	Nanchang University	1257	210	205

As can be seen from Table 1, there are 6 institutions with the number of evaluations above 200, including 2 undergraduate institutions and 4 vocational and technical colleges. The largest number of evaluations was 366, and the largest number of participants was "Web Design Fundamentals" in Beijing Language and Culture University, with 3,118 participants.

4.2. Word frequency analysis results

Frequency analysis is a research method used to analyze vocabulary by counting and analyzing the frequency of words in a given length of text, thereby revealing lexical patterns [8]. After preprocessing and word segmentation, words are counted for their frequency in the processed text, generating a frequency table. During frequency analysis, words of other parts of speech are filtered out, and only nouns, proper nouns, and gerunds are retained. The words in the frequency table are sorted from high to low, and the top 15 words are extracted for analysis.

Table 2.Word frequency analysis table of 6 courses

Course 1 Word Frequency	Course 2 Word Frequency	Course 3 Word Frequency	Course 4 Word Frequency	Course 5 Word Frequency	Course 6 Word Frequency
Teacher	Teacher	Teacher	Teacher	Teacher	Teacher
Course	Course	Course	Course	Course	Course
Content	Feelings	Content	Content	Content	Web
Details	Content	Web page	Feelings	Web page	Design
Teacher's lesson	Web	Basic	Knowledge points	Design	Content
Students	Basic	Design	Teaching	Teaching	Feel
Serious and responsible	Video	Feel	Video	Students	Website
Awesome	Design	Courseware	Web	Approach	Production
Style	Comprehensive	Platform	Comprehensive	Focus	Basics
Expectation	Production	Knowledge Points	Stick	Knowledge points	Course content
Awesome	Examples	Students	Production	Beginners	Suggestions
Code	Overall	Video	Process	Characteristics	Knowledge points
Start	Assignments	Classmates	Problems	Example problems	Theory
Feelings	Software	Questions	Software	Knowledge structure	Process
Teaching	Problems	Assignments	Time	Clearly layered	Certificate

As shown in Table 2, the top two words in the courses of the six institutions are "teacher" and "course", which shows that the comments are mainly about the teacher or the course. In Course 1, the main comments were about the content and details of the teacher's class. From the word frequency, it can be seen that most of the learners' comments about the teacher's teaching content and course were serious and responsible, awesome and beyond expectation, in which "awesome" was understood as a noun,

meaning great. In Course 2, there is more mention of specific course content, including foundational knowledge, comprehensiveness, teaching materials, teaching video quality, and learner experiences with assignments, software, and encountered problems. Course 3 primarily involves evaluations of specific course content, courseware, and knowledge points. Additionally, there are some comments related to the Chinese universities' MOOC platform. In Course 4, comments converge on the course content, teaching videos, and teaching process. The evaluations mainly focus on comprehensiveness, excellence, and also delve into detailed aspects such as knowledge points, questions, software, and time-related matters. Course 5 includes subjective assessments of teaching methods, content adaptability to learners, knowledge structure, content hierarchy, and course characteristics. It also covers more specific aspects such as knowledge points, key concepts, and example problems. In Course 6, the comments mainly concern the course content, including foundational knowledge, whether there is an excessive amount of theory, and learners' suggestions. The issue of certification is also mentioned in this course.

In summary, the reviews of the 6 courses mainly involve the evaluation of teaching contents and teachers' teaching methods, etc., of which the former is the key evaluation content. Only one of the six courses mentions certificates. Relatively speaking, the content of learners' comments is largely similar to the school's teaching evaluation, especially course 1. In a way, the more detailed and specific the learners' questions are, the more they know about the course content and the more serious they are in their independent study of the course.

4.3. Emotional Analysis

In this study, sentiment analysis was introduced to the ROSTCM6 text mining tool, and the preliminary sentiment analysis of the six courses was conducted through its own sentiment analysis tool, and the sentiment analysis results of ROSTCM6 are shown in table 3.

Table 3.ROSTCM6 sentiment analysis results graph

Course number	Positive emotions	Neutral situation	Negative mood	Course number	Positive Emotion	Neutral situation	Negative emotion
Course 1	72.14%	24.34%	3.52%	Course 4	63.79%	27.59%	8.62%
Course 2	73.77%	19.14%	7.10%	Course 5	75.20%	22.40%	2.40%
Course 3	74.19%	19.35%	6.45%	Course 6	72.41%	18.39%	9.20%

From Table 3, it can be seen that among the six courses, positive emotions accounted for the highest percentage and negative situations accounted for the least, and in general, the comments of the six courses were more positive. In the spirit of rigorous research, we checked the analysis results of ROSTCM6 and found that the overall correct rate of emotion classification is high, but there are also some misclassified data, such as "not bad, not bad", "great, great", "teacher is great, great" and so on, "The teacher is great, it's great," etc. were rated as negative emotions, "exceeded learning expectations" was defined as a neutral emotion, "too dogmatic, not very meaningful, but as a course, it doesn't seem good not to talk like this " was defined as a positive emotion. The actual positive comments should be more than what is shown in ROSTCM6.

4.4. Focused Rubric Analysis

In the fields of consumer psychology and customer management, consumer silence behavior is one of the important research topics. "Consumer silence behavior" refers to the state where consumers withhold their opinions about service problems that occur and do not file a complaint or directly express their dissatisfaction to the service organization [9]. When consumers feel dissatisfied, they may not necessarily provide feedback directly to the company. Instead, they may choose not to take any action or engage in behaviors such as sharing their experience with friends and family, or discontinuing their purchases [10]. This theory is also applicable in the context of online open courses, where a similar phenomenon of "learner silence" can occur. On the Chinese universities' MOOC platform, learners include not only working professionals and students from other schools but also a considerable number of students from the same university. For students within the same university, besides showing respect to their teachers, they may hesitate to express negative opinions due to fear of being labeled as a "problematic learner" or concerns about potential consequences of their feedback. On the other hand, learners from other institutions may choose to remain silent for various reasons and may opt to switch to other courses or directly abandon the course if they find it unsatisfactory. Additionally, some learners may provide course evaluations by simply giving a thumbs-up or liking comments that they agree with. Japanese restaurant industry expert Kazuhiko Okubo believes that one dissatisfied person implies 33

dissatisfied customers behind them. Although each industry may have different circumstances, paying attention to negative comments and highly-rated comments in course reviews can help better understand learners' needs, concerns, and identify issues with the course. These are typically areas that teachers should focus on when improving future courses. Table 4 contains the top-rated comments for each course.

Table 4. The most liked comments of the course

Course number	Highly liked content
Course 1	It smells so good!!!
Course 2	There are so many advantages, I appreciate it. Let's talk about the shortcomings. First, it's not something you can learn from a basic level. Second, some of the class content is very abrupt, the teacher directly used, but we do not know what it is. Third, some definitions are not deep enough. Fourth, the questions raised by the teacher in the class were expected to be answered at the end of the course.
Course 3	I took the course to get started with web design
Course 4	Students, thanks for supporting the Web Web Development course, which has been offered for the third time from September 24, 2019, and will be expanded to improve the first teaching, so please continue to register for the third time, and help to forward to promote the course, thanks again!
Course 5	Very good course, helps front-end development, if you are not a genius you have to study hard!
Course 6	The teacher only talked about the general, and will not how to achieve, no demonstration of the process of code. Hope to see a practical part, pure theory to understand and easy to forget

A careful analysis of the top liked comments for the courses of the six institutions revealed that the meaning of the high liked comments for course 1 was not clear; course 2 and course 6 were clear and specific comments from learners about the course, which had high value; course 3 and course 5 were messages from learners stating the purpose or benefit of the course; and course 4 was a message from the course instructor, which was embraced by the learners. In addition to the highest-liked comments, other high-liked comments are also important elements of the value analysis and will not be repeated here for reasons of space.

In addition to the high likes, the content of negative comments from learners also needs to be focused on. Table 5 shows some of the negative comments in each course.

Table 5. Some negative comments content

Course number	Some negative comments
Course 1	The content of the course is good, but there is no material for classroom cases, and it's not possible to follow the exercises, which is a pity.
Course 2	Garbage The style is not included in the style tag, but it's made directly from the dw design. If you don't talk about the code, why show the code in the PPT?
Course 3	The sound is too low, sometimes it's okay, sometimes it's not even audible when the sound is turned up, and students who live next to the highway are uncomfortable.
Course 4	It's difficult, that's all
Course 5	The questions are generally many options, but if you choose something other than the one set by the teacher, you are wrong.
Course 6	I don't understand, I feel like the teacher is reading the lesson

The analysis revealed that basically every negative comment was clear and specific, which shows from the side that every negative commenter has carefully watched the learning video. Although each learner has a different level of comprehension and knowledge, and may give negative comments that may not be fully consistent with reality. Overall, however, negative comments are something that require the course instructor's focused attention. In the analysis of the negative comments, some few major problems were found in the online open courses of web design and production:

(1) No course material, learners have no way to follow the teacher to practice.

(2) The teacher is not detailed and thorough enough, such as style tags, the difference between ID style and class style is not clearly explained, etc.

(3) The software version used is not the same, there are difficulties in operation.

(4) The course content is difficult to follow.

(5) After the course is closed, I don't know my grades, and the submitted assignments are missing.

(6) Has been abandoned tags and knowledge, I hope not to speak again, only in use, the latest.

(7) Did not have time to take the test, I do not know how to re-take the score.

(8) The video sound is too small, can not hear.

(9) Some problems with the teacher's practice questions.

(10) The teacher lectures, too much theory, I hope to combine with examples of exercises.

5. Summary and Recommendations

The above analysis shows that learners' comments are mainly about the teacher's work or the content of the course, and the comments given by learners are mainly positive, and their satisfaction with the course is relatively high. Learners may comment on the course for the sake of commenting, treating the online course comments as the teacher's comments at school, or the phenomenon of "learner silence". Therefore, teachers should focus on high likes and negative comments, and detailed and specific comments are usually closer to learners' real opinions, which are more conducive to finding problems in the course. Based on the above analysis, the following recommendations are made:

(1) In the preparatory stage of the course, the software should be prepared, and learners should be informed in advance of the version of the software to be used, the course schedule, assignment submission and other precautions on the course details page.

(2) In the course content, it should keep up with the times, follow the current trend, prepare course cases, provide materials for practice, and replace too much theoretical explanation with practical exercises.

(3) In terms of video production, attention should be paid to the duration, confirming whether the picture and sound are appropriate, and adding subtitles to avoid not knowing what the teacher is talking about and which label he/she is talking about due to accent or learners' lack of knowledge about the course.

(4) In terms of assignments and exam arrangements, learners should be reminded that in addition to paying attention to submitting assignments, they should save their assignments locally to avoid not being able to find their previous assignments after the course is closed; considering the different identities of learners, exams should be reasonably scheduled, and exam arrangements should be released in advance.

(5) Teachers can communicate with learners in the message area and answer learners' questions in a timely manner. Timely answers to learners' questions can reduce learners' sense of bewilderment and frustration when studying independently, which to some extent reduces the difficulty of learning and improves the completion rate of the course.

In addition, for programs that offer certificates, they are also sought after by students to some extent.

6. Conclusion

The data selected for this study are from web design and production courses in China's MOOC platform with more than 200 course comments. The courses include both undergraduate and vocational schools. A combined qualitative and quantitative analysis approach was used. The results from word frequency analysis, sentiment analysis, and analysis of highly-rated and negative comments indicated that word frequency analysis can reveal the main content of course comments, with learners' positive comments significantly outweighing the negative ones. Analyzing highly-rated and negative comments can accurately identify learners' opinions and needs regarding the course, which is a focal point for teachers. The findings of this study provide references for the development of online open course platforms, improvement of online course design and organization. However, this study adopted a relatively brief approach in topic analysis. Subsequent research could incorporate TF-IDF and LDA methods to discover more specific themes in the comments. Additionally, this study only focused on courses from the Chinese MOOC platform. Future research could include comparative analysis across multiple platforms and comparative analysis among different learner types to identify learners' focal points and explore their learning motivations.

Acknowledgements

References

[1] Adams A A, Williams S A. MOOCs: A systematic study of the published literature 2008-2012[J]. International review of research in open and distributed learning, 2013, 14(3): 202-227.

[2] Wu, Huajun, He, Juhou, Chen, Qitie et al. Sentiment Analysis and Topic Mining for Online Excellence Course Evaluation of Vocational Education[J]. China Vocational and Technical Education, 2022,No.798(02):55-63.

[3] Pan Fang,Zhang Huibing,Dong Junchao et al. Efficient Transformer-based sentiment analysis in terms of Chinese online course reviews[J]. Computer Science,2021,48(S1):264-269.

[4] Hew K F, Hu X, Qiao C, et al. What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach[J]. Computers & Education, 2020, 145: 103724.

[5] Ye D, Pennisi S. Analysing interactions in online discussions through social network analysis[J]. Journal of Computer Assisted Learning, 2022, 38(3): 784-796.

[6] Wen M, Yang D, Rose C. Sentiment Analysis in MOOC Discussion Forums: What does it tell us?[C]//Educational data mining 2014. 2014.

[7] Hew K F, Hu X, Qiao C, et al. What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach[J]. Computers & Education, 2020, 145: 103724.

[8] Luo Yan,Zhao Shuliang,Li Xiaochao et al. A method of text keyword extraction based on word frequency statistics[J]. Computer Applications,2016,36(03):718-725.

[9] Wang WF,Xu ZH. Consumer silent behavior in service failure context and its impact on brand relationship[J]. Journal of Jinan (Philosophy and Social Science Edition),2015,37(06):42-51.

[10] Singh J. Consumer complaint intentions and behavior: definitional and taxonomical issues[J]. Journal of marketing, 1988, 52(1): 93-107.