



# Leveraging text mining and analytic hierarchy process for the automatic evaluation of online courses

Xieling Chen<sup>1</sup> · Haoran Xie<sup>2</sup> · Xiaohui Tao<sup>3</sup> · Fu Lee Wang<sup>4</sup> · Jie Cao<sup>5</sup>

Received: 8 October 2023 / Accepted: 3 May 2024 / Published online: 21 May 2024  
© The Author(s) 2024

## Abstract

This study introduced a multi-criteria decision-making methodology leveraging text mining and analytic hierarchy process (AHP) for online course quality evaluation based on students' feedback texts. First, a hierarchical structure of online course evaluation criteria was formulated by integrating topics (sub-criteria) identified through topic modeling and interpreted based on transactional distance and technology acceptance theories. Second, the weights of the criteria in the hierarchical structure were determined based on topic proportions. Third, the AHP was employed to determine the overall relative advantage of online courses and their relative advantage within each criterion based on the hierarchical framework and criterion weights. The proposed approach was implemented on the datasets of 6940 reviews for knowledge-seeking courses in Art, Design, and Humanities (D1) and 44,697 reviews for skill-seeking courses in Computer Science, Engineering, and Programming (D2) from Class Central to determine ranking positions of nine courses from both D1 and D2 as alternatives. Results revealed common concerns among knowledge and skill-seeking course learners, encompassing “assessment”, “content”, “effort”, “usefulness”, “enjoyment”, “faculty”, “interaction”, and “structure”. The article provides valuable insights into the online course evaluation and selection processes for learners in D1 and D2 groups. Notably, both groups prioritize “effort” and “faculty”, while D2 learners value “assessment” and “enjoyment”, and D1 learners value “usefulness” more. This study demonstrates the efficacy of leveraging online learner reviews and topic modeling for automating MOOC evaluation and informing learners' decision-making processes.

**Keywords** Topic mining · Analytic hierarchy process (AHP) · Automatic evaluation · MOOCs · Course selection

✉ Fu Lee Wang  
pwang@hkmu.edu.hk

Xieling Chen  
xielingchen0708@gmail.com

Haoran Xie  
hrxie2@gmail.com

Xiaohui Tao  
Xiaohui.Tao@usq.edu.au

Jie Cao  
cao\_jie@hfut.edu.cn

<sup>1</sup> School of Education, Guangzhou University, Guangzhou, China

<sup>2</sup> School of Data Science, Lingnan University, Tuen Mun, Hong Kong SAR, China

<sup>3</sup> School of Mathematics, Physics, and Computing, University of Southern Queensland, Springfield, Australia

<sup>4</sup> School of Science and Technology, Hong Kong Metropolitan University, Ho Man Tin, Hong Kong SAR, China

<sup>5</sup> School of Management, Hefei University of Technology, Hefei, China

## 1 Introduction

Massive open online courses (MOOCs), enabling the delivery of high-quality education on an unprecedented level in terms of cost-effectiveness and worldwide accessibility [1], have captured the attention of researchers to explore diverse topics such as MOOC classifications, learning engagement, and concept recommendations. For example, utilizing 102,184 reviews across 401 MOOCs as the foundation, Chen et al. [2] devised DNN-powered models to autonomously differentiate a set of semantic groupings. Wei et al. [3] examined the correlations between motivation, perceived support, engagement, and self-regulated learning strategies concerning learners' perceived outcomes in MOOCs using an online survey involving 546 participants. Gong et al. [4] focused on reinforcement learning and heterogeneous information network-driven concept recommendations by leveraging the interactions between users and knowledge concepts, and among users, courses, videos, and concepts. Gong et al. [5]

proposed an attention-driven, heterogeneous graph convolutional deep knowledge recommender designed to suggest knowledge concepts within MOOCs. The recommender harnessed content and contextual information to master entity representation through graph convolution networks.

The proliferation of MOOCs has generated concerns about and prompted extensive research on courses' quality and effectiveness. Currently, MOOC websites only provide rough overall ratings, making it challenging to differentiate between various courses, especially in the cases of overall course ratings displayed on the MOOC website. Therefore, there is a pressing need to obtain detailed insights into the performance of MOOCs based on different evaluation criteria [6]. As learners have diverse needs, some may prioritize assessment aspects, while others may value interaction. By displaying course performance across different dimensions, learners can identify courses that excel in criteria that are most relevant to their preferences.

Traditional top-down methods for MOOC course evaluation included surveys, expert interviews, and literature reviews [7–9]. Survey-based methods are frequently utilized for exploring learner engagement, satisfaction, and the intention to continue with courses. For instance, Dai et al. [9] revealed that learners' attitude significantly impacted their intent to continue, and MOOC instructors should be cautious in their course promotions to avoid overemphasizing benefits. Similarly, based on 622 structured questionnaires from undergraduate students in Malaysia, Albelbisi et al. [8] revealed that (1) system quality positively influenced satisfaction, (2) satisfaction and service quality positively affected self-regulated learning, and (3) system quality affected self-regulated learning through satisfaction. Interviews, although involving fewer participants, present the advantage of flexible and in-depth questioning to gain deeper insights into learners' motivations, course completion, and satisfaction levels. For instance, Zhu et al. [7] carried out semi-structured discussions with 15 online learners, revealing that learners' contentment was influenced by course design factors (such as well-structured organization) and instructional methodologies (such as instruction presence). However, surveys and interviews suffer from delayed feedback, time and cost constraints, expert-centricity, and a lack of learner perspectives. To address these limitations, adopting a scientific approach for systematically and objectively assessing the quality of online courses becomes crucial. Such an approach is pivotal in diagnosing and improving online course quality.

Due to the progress in big data and text mining methodologies, scholars have shifted their focus towards utilizing online course review data to obtain valuable insights for enhancing course quality [10]. This approach facilitates a more comprehensive comprehension of learners' emotions, actions, course acceptance, platform comparisons,

and prevailing trends, contributing to a better grasp of the crucial elements influencing MOOC success [11]. However, existing methods for determining online course quality evaluation indicators and [or] their weights often rely on group decision-making [1, 12, 13], resulting in poor adaptability and applicability for fine-grained evaluation of different types of courses. Expert simulation evaluations and the use of pre-determined indicators may not accurately reflect learners' experiences and requirements, resulting in a lack of learner-centeredness in course evaluation. Thus, there is a need to leverage crowdsourced data from MOOC platforms, artificial intelligence (AI), and text analysis techniques to empower researchers to effectively tap into the collective wisdom and expertise of learners to enhance course design and learner contentment [14].

To address deficiencies in prior studies, this study focuses on learners' learning experiences and needs, breaking away from the reliance on experiential judgment and excessive human subjectivity in traditional online course quality research. The objective is to leverage unstructured data from student feedback texts, utilizing text mining and hierarchical structure modeling to develop an intelligent evaluation model and implementation plan for online course quality.

Accordingly, the present study formulates three research questions (RQs):

*RQ1:* What factors will be included in the multi-criteria decision-making framework? What are the similarities and differences between the selection of knowledge- and skill-seeking courses?

*RQ2:* What are the most influential factors in course selection? What are the similarities and differences between the selection of knowledge- and skill-seeking courses?

*RQ3:* What are the ratings of courses and how do they rank according to sub-criteria and criteria?

This research endeavors to answer these questions by presenting a novel crowdsourcing technique that utilizes text mining and analytic hierarchy process (AHP) to automatically evaluate MOOCs. The evaluation is based on the aggregation of learners' reviews collected from 169 MOOCs on Class Central. Considering RQ1, this study integrates topics (sub-criteria) identified through topic modeling and interpreted under the framework of transactional distance and technology acceptance theories to form the hierarchical structure of MOOC evaluation criteria. Regarding RQ2, this study leverages the probability distribution of topics (sub-criteria) identified through topic modeling to weigh the relative importance between criteria. Subsequently, addressing RQ3, based on the established hierarchical structure and the respective criterion weights, this study employs AHP to rank the online courses to

determine their overall relative advantage and relative advantage within each criterion. By doing so, this study can provide a finely-tuned analysis methodology for large-scale online educational course quality evaluation using a text mining perspective, thus providing the necessary technical support for investigating large-scale online course quality evaluation.

## 2 MOOC evaluation based on review mining

Currently, online course quality evaluation research primarily focuses on the construction of evaluation indicators and models. These studies often rely on traditional methods such as literature review, questionnaire surveys, and expert scoring. However, these methods are time-consuming and costly, and they are subject to experiential evaluations and subjective interventions, resulting in poor adaptability of indicators and difficulty in conducting detailed evaluations for different course types [12, 15].

Consequently, the investigation into developing theories to assess the caliber of online courses using text mining from student feedback has garnered considerable interest. The main achievements in this field involve using topic modeling to automatically identify latent evaluation topics from text data [16–18]. Through the amalgamation of findings from topic modeling and undertaking theoretical examination, course evaluation indicators can be formulated. Based on this, course evaluation indicators can be hierarchically aggregated according to their affiliations, forming a multi-level analytical structural model, thus accomplishing the establishment of the theoretical framework for course evaluation, and authentically reflecting the learners' demands, guiding the course evaluation practice effectively.

Traditional research on course quality evaluation often relies on qualitative group decision-making and expert simulation evaluation methods, neglecting the role of learners as the primary stakeholders. Moreover, the use of pre-determined indicator factors often results in measurement items that fail to accurately reflect learners' learning experiences and needs, making it difficult to achieve large-scale, normalized, and continuous course quality monitoring. However, learners' perceived learning experiences are crucial references for online course design and quality improvement. Therefore, it is necessary to focus on learners' experiences and needs and use text mining in course evaluations to provide important foundations for the evaluation of online teaching effectiveness from the learners' perspective [1].

However, existing research on online course quality evaluation using text mining often relies on subjective assignment methods such as AHP [19] to obtain indicator weights [1], heavily relying on expert opinions and

judgments, which fail to reflect the learners' true experiences and needs. By utilizing the estimated popularity of evaluation topics obtained via topic modeling as weights for course evaluation indicators to represent the learners' levels of interest, achieving automated customization of indicator weights and enabling automatic sorting of various courses to be evaluated [20].

## 3 Data preparation

### 3.1 Review dataset collection

The MOOC reviews were scraped from Class Central using a self-developed crawler, which was then parsed into Excel files for further processing. Upon removing duplicates and MOOCs with less than 20 review comments, non-English reviews were identified and subsequently excluded. Using a Python package named “langid”, a standalone language identification tool. Python with TextBlob was used to automatically check and correct spelling. For instance, “I havv goood speling” was corrected to “I have good spelling”. Two types of courses were considered in this study. The first relates to Art, Design, and Humanities, which is generally in a domain that is “knowledge-seeking”. The second relates to Computer Science, Engineering, and Programming, which is generally in a domain that is “skill-seeking”. As a result, a total of 52,881 reviews were obtained.

### 3.2 Helpful review identification

Among the 52,881 reviews, 4407 reviews had a helpful vote value. For these 4407 reviews, this study followed O'Mahony and Smyth [21] to define the top 75% ranked by the number of the helpful votes they received as helpful reviews, while the rest were treated as unhelpful reviews. An examination of 50 randomly selected unhelpful reviews suggested that they mostly lacked valuable information pertaining to specific aspects of MOOC courses. Exemplified included: “great class”, “the class was amazing”, “omg so good”, “I loved this class”, “one of the best I have ever taken”, “strongly recommended”, and “I don't know what to say”. As a result, 3305 helpful reviews, labeled as “1” and 1102 unhelpful reviews, labeled as “0”, were randomly divided into training and testing datasets to train and test the classifier based on a Naive Bayes model and Word Level TF-IDF to automatically predict labels (“1” and “0”) of the 48,474 reviews without a helpful vote value. The prediction results, combined with the previously labeled sample, comprised the datasets of D1 and D2 for (1) Art, Design, and Humanity and (2) Computer Science, Engineering, and Programming courses, respectively. The numbers of courses and reviews for D1 and D2 datasets are presented in Table 1. Specifically, the D1 dataset comprised 63 courses

**Table 1** Number of courses and reviews for the D1 and D2 datasets

Datasets	Subjects	Number of courses		Number of reviews	
		Number	Total	Number	Total
D1	Art and design	12	63 (37.28%)	376	6940 (13.44%)
	Humanities	51		6564	
D2	Computer Science	45	106 (62.72%)	1852	44,697 (86.56%)
	Engineering	23		1501	
	Programming	38		41,344	

and 6940 helpful reviews, taking proportions of 37.28% and 13.44% of the total number of courses and helpful reviews included, respectively. The D2 dataset comprised 106 courses and 44,697 helpful reviews, taking proportions of 62.72% and 86.56% of the total number of courses and helpful reviews included, respectively.

### 3.3 Data preprocessing

To preprocess the data for topic modeling, six sequential actions were executed: (1) Tokenization and exclusion of special characters, (2) normalization, (3) elimination of stop words, (4) lemmatization, (5) term selection, and (6) the formation of a term-document matrix. In step 1, every sentence extracted from review comments was disassembled into a word list, with the exclusion of special characters and punctuation. Step 2 converted uppercase letters to lowercase. Step 3 played a pivotal role in the data preprocessing process. During this phase, numbers, punctuations, symbols, and stop-words (e.g., “me”, “I”, “or”, “him”, “a”, and “they”) were excluded, as they “appear frequently and are insufficiently specific to represent document content (p. 976)” [22]. Step 4 aimed to analyze vocabulary and morphological aspects of words. Its primary objective is to eliminate inflectional endings and obtain the root form. For instance, the lemma of “courses” is “course”, and “assessing” is transformed into “assess”, while “mice” becomes “mouse”. This lemmatization process holds great importance in text mining. Another common method for reducing terms to their base form is stemming. However, stemming was not utilized as it often collapses derivationally related words [23]. For instance, the stem of “organized” would be “organ”, while the lemma remains “organize”. In this case, stemming may lead to challenges in correctly interpreting term stems. After lemmatization, less significant terms were excluded in step 5 using TF-IDF. In step 6, we created a term-document matrix, representing the occurrence of terms (rows)

in documents (columns), thereby constituting the corpus for the STM algorithm.

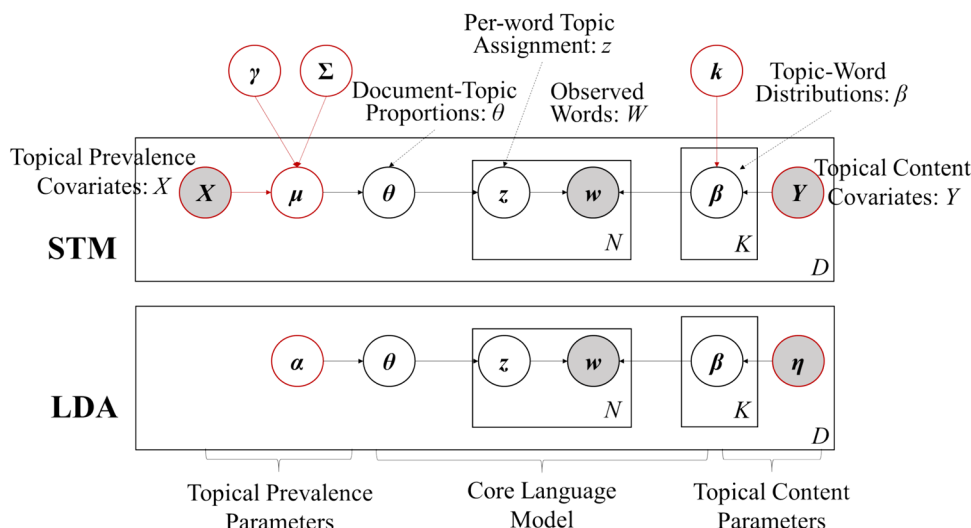
## 4 Methods

### 4.1 Creating STM models

In our analysis of MOOC reviews, we adopted the structural topic model (STM) [24] as it offers notable advancements compared to latent Dirichlet allocation (LDA). LDA and STM are both Bayesian generative frameworks employed in the field of topic modeling. They posit that a topic embodies a probability distribution over words, and each document comprises a blend of topics spanning the entire corpus [25]. However, STM presents significant enhancements by incorporating document-level structural information, which influences the prevalence of topics at the document level (i.e., per-document topic proportions) and the distribution of words within a topic (i.e., topic-word distributions). This approach allows for a more targeted examination of the influence of covariates on the text content.

Figure 1 presents a graphical representation of the technical distinctions between STM and LDA. Each variable is depicted as a node. The nodes without shading indicate latent variables, whereas the nodes with shading denote observable variables. The rectangular shapes denote replication:  $n \in \{1, 2, \dots, N\}$  signifies the indexing of words within a document;  $k \in \{1, 2, \dots, K\}$  is the indexing of each topic, based on a predefined number of topics,  $K$ ; and  $d \in \{1, 2, \dots, D\}$  denotes the indexing of documents. Additionally, Fig. 1 depicts that solely node  $w$  (i.e., words within documents) is directly measurable in the two models. Therefore, the primary objective lies in deducing the concealed topic information from the observable words,

**Fig. 1** Plate diagram comparison of LDA and STM adapted from [26]



$W$ , and producing matrices: document-topic proportions,  $\theta$ , and topic-word distributions,  $\beta$ .

In this study, the textual information from the beneficial reviews in D1 and D2 served as the data input for topic modeling. For both datasets, we generated 36 distinct STM by altering the  $K$  within the range of 5 to 50.

### 4.2 Exclusivity and coherence measures

This study selected models by considering exclusivity and semantic coherence metrics, alongside manual verification. Mimno et al. [27] introduced semantic coherence as a measure closely associated with pointwise mutual information [28]. This criterion reaches its peak when the most likely words within a particular topic frequently appear together. Mimno et al. demonstrated a strong correlation between semantic coherence and human evaluations of topic quality. To formalize this, consider  $D(v_i, v_j)$  as the count of occurrences where words  $v_i$  and  $v_j$  coexist within a document. The semantic coherence for topic  $k$  is given by Roberts et al. [29] as Eq. (1), where  $M$  represents the top  $M$  most likely words within topic  $k$ . Each model computes an aggregate coherence score by determining the coherence of each topic separately and subsequently averaging these individual values.

$$C_k = \sum_{i=2}^M \sum_{j=1}^{i-1} \log \left( \frac{D(v_i, v_j) + 1}{D(v_j)} \right) \tag{1}$$

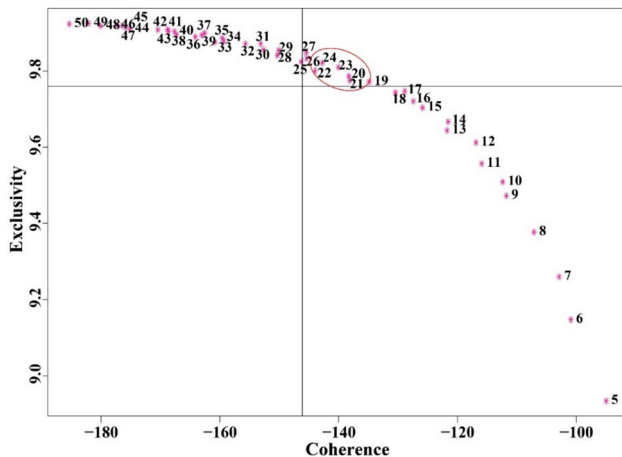
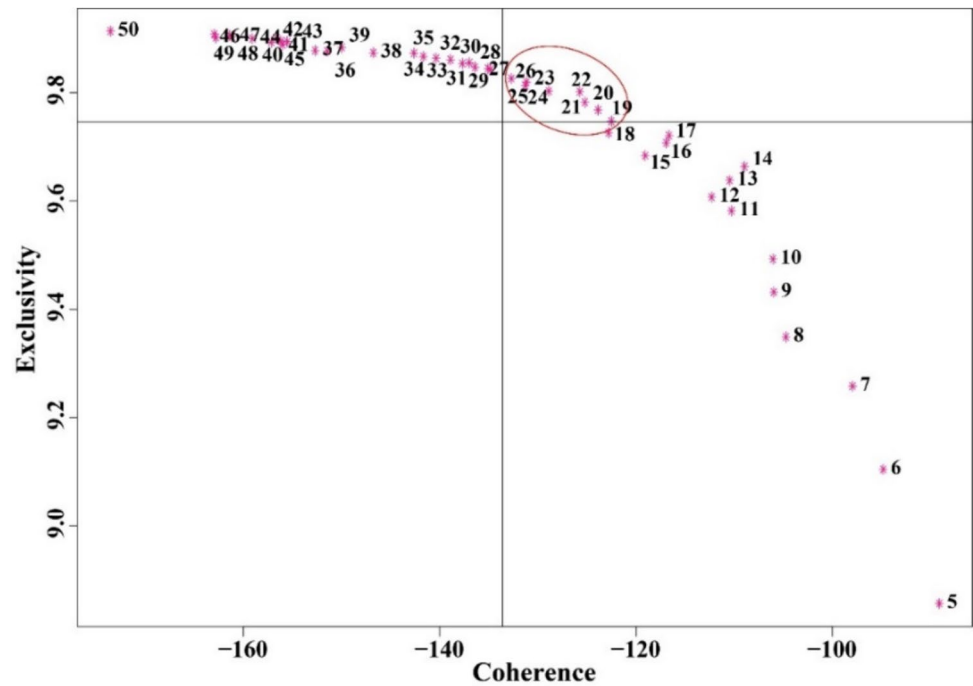
Exclusivity measures the degree to which the primary words within a topic are unique to that specific topic and not prevalent among others. This value is essentially an average, calculated for each top word, where the

probability of that word within the topic is divided by the total probabilities of that word across all topics. The FREX metric [30] quantifies exclusivity while considering word frequency. FREX represents the weighted harmonic mean of a word’s rank concerning exclusivity and frequency [29] shown as Eq. (2). ECDF stands for the empirical cumulative distribution function,  $\omega$  is the weight (typically set to 0.7 to prioritize exclusivity),  $k \in K$  denotes the  $k$ th topic,  $v$  signifies the word being evaluated, and  $\beta$  refers to the topic word distribution for that specific topic. The cumulative distribution function of a real-valued random variable  $X$ , computed at  $x$ , indicates the probability of  $X$  assuming a value less than or equal to  $x$ . In contrast, the ECDF represents the probability distribution derived from the sampled dataset rather than the entire population.

$$\text{FREX}_{k,v} = \left( \frac{\omega}{\text{ECDF}(\beta_{k,v} / \sum_{j=1}^K \beta_{j,v})} + \frac{1 - \omega}{\text{ECDF}(\beta_{k,v})} \right)^{-1} \tag{2}$$

In this study, both the coherence and exclusivity of each topic within a model are computed using the *manyTopics* function within the *stm* R package [29]. These values were then averaged across all topics to derive the model’s overall score. Models exhibiting elevated exclusivity and semantic coherence were typically favored. For D1, among the 36 estimated models, six ( $K=19, 20, 21, 22, 23, 24, 25$ , and 26) surpassed the others and were chosen for further comparisons. Subsequently, a qualitative assessment of the eight models was performed by examining terms and reviews to identify the most suitable model, which turned out to be the one with 22 topics. Hence, based on both quantitative and qualitative evaluations, the model with 22 topics was selected as the ultimate choice.

**Fig. 2** Model diagnosis based on exclusivity and coherence for the D1 dataset



**Fig. 3** Model diagnosis based on exclusivity and coherence for the D2 dataset

Employing a similar approach, the optimal topic model for D2 was found to be the one with 24 topics. Figures 2 and 3 present the outcomes of model diagnosis using exclusivity and coherence for D1 and D2.

### 4.3 Topic representation and assignment

The findings from the optimal models on D1 and D2 were used to identify the discriminant terms, also known as commonly employed words, organized in a descending sequence based on their frequency within each topic. For every topic, the topic proportions were computed

using  $P_k = (\sum_d \theta_{d,k})/D$ , based on a matrix indicating the connection between course reviews and topics in terms of proportion. In this equation,  $P_k$  represents the frequency of the  $k_{th}$  topic within the data corpus;  $\theta_{d,k}$  represents the prevalence of the  $k_{th}$  topic in the  $d_{th}$  course review, and  $D$  indicates the total number of course reviews.

To provide a more concise view of the topics, we needed to condense and categorize them. As indicated in Chang et al. [31], accomplishing this task necessitates human judgments and interventions to assign appropriate labels or titles to the topics. This is determined by the semantic relatedness of important terms within the topic matrix, ensuring a more accurate representation of the topic contents. In the quest to capture shared opinions in review comments, [32] identified six coding categories derived from transactional distance theory to grasp learners' encounters in MOOCs, consisting of "structure", "videos", "instructors", "content", "interaction", and "assessment". In line with technology acceptance theories, Du and Gao [33] developed four primary factors, encompassing usefulness, enjoyment, technicality, and effort, as potential indicators of AI-based educational application adoption. Grounded on previous studies, this study included nine factors (i.e., "assessment", "content", "effort", "enjoyment", "faculty", "interaction", "structure", "technicality", and "usefulness") as predetermined labels (criteria) to understand MOOC learners' concerns. Subsequently, based on the nine criteria, along with the semantic similarity of crucial terms within the topics, two experienced experts specializing in MOOC

education deduced and condensed sub-labels (sub-criteria) for each topic in the optimal models for D1 and D2. This labeling process adhered rigorously to a procedure widely adopted in prior topic modeling studies (e.g., [34, 35]). To maximize precision, this study also involved multiple rounds of in-depth discussions among the two experts until a consensus was reached. Topics that merely described the course content without offering any opinions or evaluations on its quality were omitted. To maintain consistency, the proportions ( $P_k$ ) of the included topics were adjusted by normalizing their entries and dividing each entry by the sum of all entries. Finally, 13 relevant topics were retained for D1, and 18 relevant topics were retained for D2.

As a result, the nine criteria, along with their sub-criteria, were utilized to highlight learners' major concerns. For example, this study introduced the predetermined label of "faculty" to encompass a comprehensive range of factors directly related to instructors within the MOOC context. The sub-labels under "faculty" encapsulate various aspects related to instructors, for example, presentation style, passion, humor, teaching methodologies, overall experience, and the collaborative dynamics within the faculty team. This study also introduced the predetermined label of "technicality" to encapsulate various essential components shaping the quality of MOOC courses. This label was structured with sub-labels, namely "complexity" and "flexibility", each of which encompasses intricate aspects that intricately involve considerations regarding the role and impact of technological factors such as videos within MOOCs [33]. Specifically, within the "complexity" sub-label, we incorporated technical aspects that are central to the quality of video content in MOOCs, for example, video resolution, streaming capabilities, interactivity, and accessibility features. Furthermore, the "flexibility" sub-label was designed to address the adaptability and user-centric features of video content within MOOCs, for example, the ability of videos to cater to diverse learning styles and their responsiveness across devices.

- "assessment": assignment, autograder
- "content": problem-solving, explanation, use of cases/examples
- "effort": perceived difficulty, perceived workload
- "enjoyment": pleasure, satisfaction
- "faculty": instructor humor, instructor presentation, instructor passion, teaching style, instructor experience, faculty team
- "interaction": interaction
- "structure": course description
- "technicality": complexity, flexibility
- "usefulness": knowledge enhancement, job preparedness, beginner friendliness

#### 4.4 Determining criteria, weights, and alternatives in AHP

This stage holds immense significance when facing a MOOC selection dilemma. Utilizing the outcomes of topic modeling for D1 and D2, we ascertain the criteria, sub-criteria, and their respective weights. Based on this information, two decision hierarchies can be created, as exemplified in the selection of MOOC courses for Art, Design, and Humanities as well as Computer Science, Engineering, and Programming, respectively. The weight assigned to each criterion is the total of the weights allocated to its sub-criteria (i.e., topic label) using the formula  $P_{cr} = \sum_{k=1} P_k$ .

In this research, only MOOC courses with over 100 reviews and an overall rating score of  $\geq 4.7$  stars (where 1 signifies terrible and 5 denotes excellent) were taken into consideration during the course selection process. This study chose to analyze courses with an overall rating score of 4.7 or above because of the common challenge users encounter when attempting to choose among courses with marginal differences in overall ratings. Our method aims to encourage users to delve deeper into the evaluation process for the initially selected highly rated courses. Thus, we specifically focused on comparing highly-rated courses across

**Table 2** Appointed alternatives for Art, Design, and Humanity courses

Labels	Course names	Overall ratings	Number of reviews
A1	A Life of Happiness and Fulfillment	4.8	375
A2	Academic Writing	4.7	1021
A3	English for the Workplace	4.8	922
A4	English in Early Childhood: Language Learning and Development	4.8	1046
A5	Exploring English: Language and Culture	4.7	628
A6	Exploring English: Shakespeare	4.8	386
A7	HOPE: Human Odyssey to Political Existentialism	4.9	177
A8	How to Succeed at: Writing Applications	4.7	282
A9	Teaching English: How to Plan a Great Lesson	4.8	271

**Table 3** Appointed alternatives for Computer Science, Engineering, and Programming courses

Labels	Course names	Overall ratings	Number of reviews
B1	Computing in Python I: Fundamentals and Procedural Programming	4.8	138
B2	Differential Equations for Engineers	4.9	165
B3	Elements of AI	4.8	552
B4	Introduction to Engineering Mechanics	4.7	148
B5	Introduction to Programming with MATLAB	4.8	147
B6	Machine Learning	4.7	132
B7	Matrix Algebra for Engineers	4.8	255
B8	Python for Everybody	4.9	312
B9	Unlocking Information Security I: From Cryptography to Buffer Overflows	4.8	159

fine-grained dimensions. Also, considering the caliber of the reviews and the extensive data at hand for assessing these courses we opted for courses with a score of  $\geq 4.7$  stars that had garnered at least 100 reviews. This method provides adequate time for assessing the course's performance. Accordingly, 9 courses relating to Art, Design, and Humanity as well as Computer Science, Engineering, and Programming were used as alternatives in the AHP (Tables 2 and 3).

#### 4.5 Course ranking via AHP

Once the hierarchical structure for MOOC selection was established, the prioritization process determined the comparative importance of the MOOCs. All alternatives undergo evaluation by learners from Class Central, making this a collective decision-making analysis.

For each course, the score assigned by a reviewer (i.e., learner) for the  $k_{th}$  topic (sub-criteria) is calculated using  $a_k = \sum_{c=1} r_c p_{ck}$ , where  $r_c$  represents the overall rating score provided by the  $c_{th}$  reviewer (i.e., learner), and  $p_{ck}$  denotes the percentage of the  $c_{th}$  review comment assigned to the  $k_{th}$  topic. Consequently, the score for each criterion is the summation of the calculated scores for each of its sub-criteria using  $a_{cr} = \sum_{k=1} a_k$ . Subsequently, the pairwise comparison matrices were formulated as Eq. (3) based on the acquired data. The pairwise comparisons consist of  $n \times n$  elements, where  $n$  denotes the number of factors under consideration.

$$a_{ij} = \frac{1}{a_{ji}}, \quad \text{where } i, j = 1, 2, 3, \dots, n \quad (3)$$

Using a nine-point scale, each element denotes the assessment of the comparative significance of two factors. Element  $a_{ij}$  signifies the extent to which the  $i$ th factor dominates the  $j$ th factor, resulting in the placement of the  $a_{ij}$  factor in the  $i$ th row and  $j$ th column, while its reciprocal is placed in the  $j$ th row and  $i$ th column. Thus, a comparison matrix is drawn

from Eq. (4). To derive priorities for each criterion, denoted as  $priorities_{cr}$ , we initially compute the principal right eigenvector for each matrix and subsequently normalize its entries through division by their collective total. Afterward, the priorities for each MOOC course can be determined by computing the weighted sum score using  $\sum(P_{cr} \times priorities_{cr})$ .

$$A = \begin{pmatrix} a_{11} & 1/a_{21} & \dots & 1/a_{n1} \\ a_{21} & a_{22} & \dots & 1/a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (4)$$

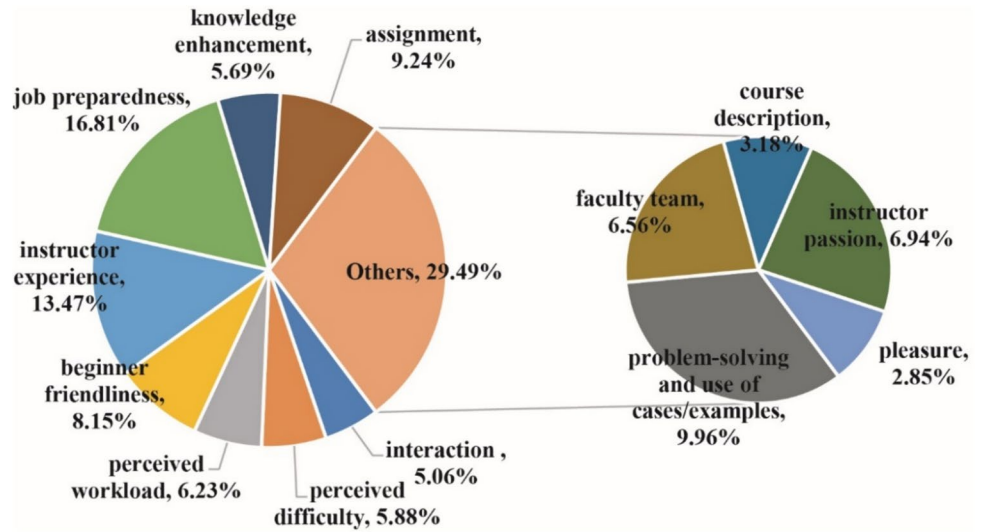
## 5 Results

### 5.1 Topic identification

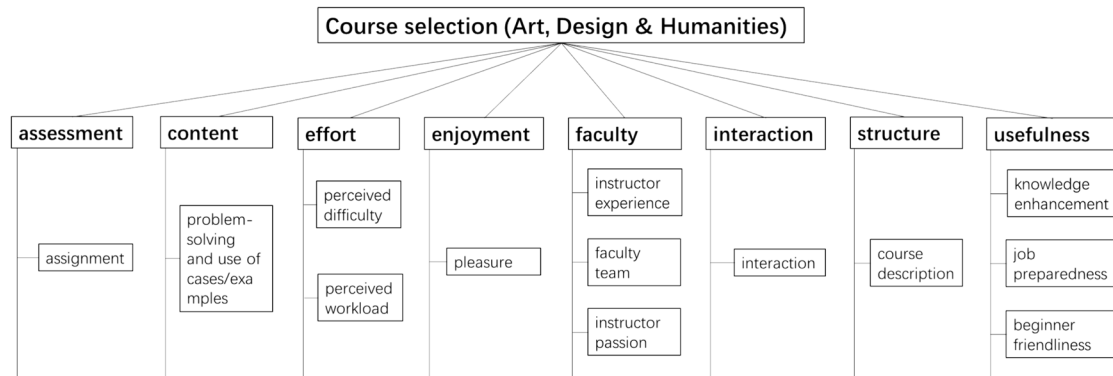
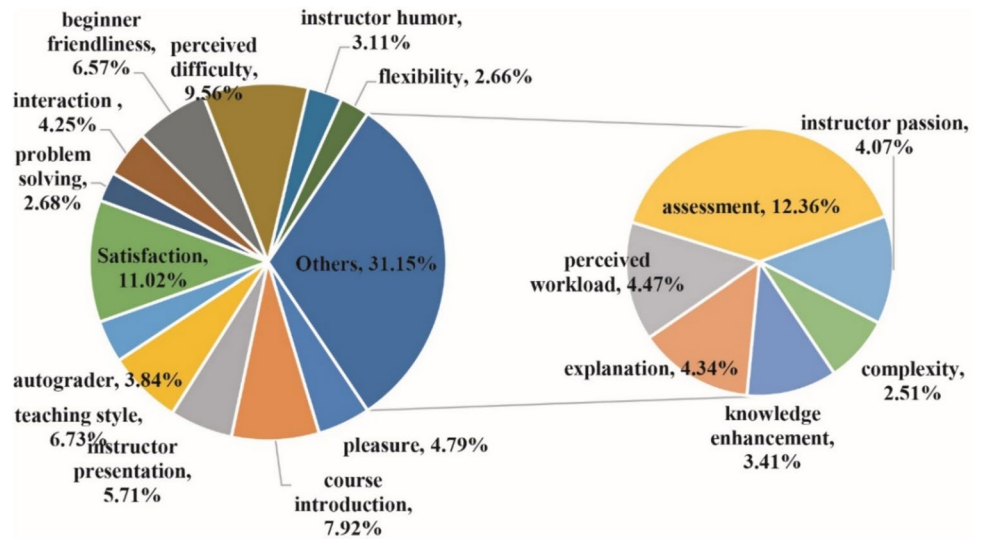
The outcomes of the optimal STM model on D1 and D2 can be found in Figs. 4 and 5 and Tables 10 and 11 in the "Appendix". Extracted from the topic modeling results for D1 and D2 are the criteria, sub-criteria (referred to as effective topic labels), and adjusted weights (termed adjusted topic proportions). The 8 criteria considered in selecting Art, Design, and Humanities courses are "assessment", "content", "effort", "enjoyment", "faculty", "interaction", "structure", and "usefulness". Furthermore, some criteria have sub-criteria. For instance, the criterion of "usefulness" encompasses sub-criteria of "knowledge enhancement", "job preparedness", and "beginner friendliness". As for the selection of Computer Science, Engineering, and Programming courses, 9 criteria were considered: "assessment", "content", "effort", "enjoyment", "faculty", "interaction", "structure", "technicality", and "usefulness". Similar to the previous case, certain criteria possess sub-criteria at the subsequent level. For instance, the criterion of "faculty" consists of



**Fig. 4** Extracted valid topics and their adjusted weights for D1



**Fig. 5** Extracted valid topics and their adjusted weights for D2



**Fig. 6** Decision hierarchy for Art, Design, and Humanities course selection

sub-criteria of “instructor humor”, “instructor presentation”, “instructor passion”, and “teaching style”.

Using the identified criteria and sub-criteria, two decision hierarchies were formulated, as illustrated in Figs. 6 and 7, for MOOC selection. Next, by referring to the adjusted weights (adjusted topic proportions) presented in Fig. 4, the weight of the criterion “usefulness” is computed as follows: 0.0569 (adjusted weight of “knowledge enhancement”) + 0.1681 (adjusted weight of “job preparedness”) + 0.0815 (adjusted weight of “beginner friendliness”) = 0.3065. For criteria without any sub-criteria, such as “interaction”, the adjusted weight of “interaction”,

which is 0.0506, is used directly. The weights for other criteria are as follows: 0.1211 (“effort”), 0.0924 (“assessment”), 0.0996 (“content”), 0.2697 (“faculty”), 0.0318 (“structure”), and 0.0285 (“enjoyment”).

Similarly, by examining the adjusted weights (adjusted topic proportions) presented in Fig. 5, the weights of the criteria are as follows: 0.1581 (“enjoyment”), 0.0792 (“structure”), 0.1963 (“faculty”), 0.1619 (“assessment”), 0.0703 (“content”), 0.0425 (“interaction”), 0.0998 (“usefulness”), 0.1403 (“effort”), and 0.0517 (“technicality”).

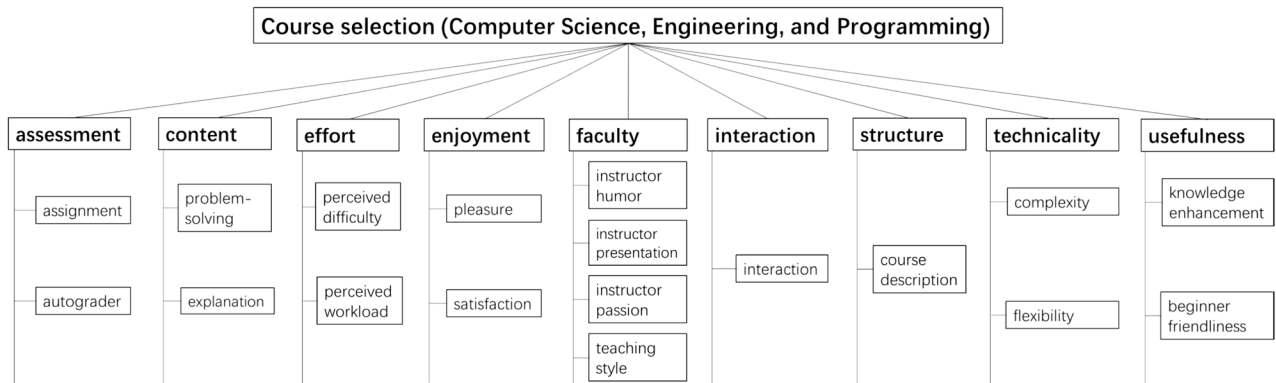


Fig. 7 Decision hierarchy for Computer Science, Engineering, and Programming course selection

Table 4 Scores for each criterion in Art, Design, and Humanity course selection

Labels	Assessment	Content	Effort	Enjoyment	Faculty	Interaction	Structure	Usefulness
A1	93.058	69.998	90.067	30.566	786.788	48.771	20.047	148.929
A2	170.572	270.415	341.281	71.524	1221.083	197.756	66.295	481.374
A3	233.650	371.206	425.448	80.238	488.463	107.347	37.935	1486.218
A4	141.120	289.192	298.937	89.750	754.454	94.699	39.696	635.172
A5	145.021	177.206	152.653	53.380	295.626	73.850	32.373	545.128
A6	71.500	83.428	153.605	34.029	199.261	52.621	18.353	223.457
A7	30.051	44.532	45.300	16.041	112.271	31.747	53.369	91.376
A8	39.190	101.427	78.868	23.492	154.464	36.637	18.587	297.009
A9	41.202	102.389	76.182	21.913	469.297	27.531	14.453	200.820

Table 5 Pairwise comparison matrix for the alternatives regarding assessment in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.546	0.398	0.659	0.642	1.302	3.097	2.375	2.259	0.08041060	5
A2	1.833	1	0.730	1.209	1.176	2.386	5.676	4.352	4.140	0.14738911	2
A3	2.511	1.370	1	1.656	1.611	3.268	7.775	5.962	5.671	0.20189425	1
A4	1.516	0.827	0.604	1	0.973	1.974	4.696	3.601	3.425	0.12193964	4
A5	1.558	0.850	0.621	1.028	1	2.028	4.826	3.700	3.520	0.12531078	3
A6	0.768	0.419	0.306	0.507	0.493	1	2.379	1.824	1.735	0.06178207	6
A7	0.323	0.176	0.129	0.213	0.207	0.420	1	0.767	0.729	0.02596659	9
A8	0.421	0.230	0.168	0.278	0.270	0.548	1.304	1	0.951	0.03386392	8
A9	0.443	0.242	0.176	0.292	0.284	0.576	1.371	1.051	1	0.03560174	7

**Table 6** Numerical values for ratings of the alternatives in Art, Design, and Humanity course selection

Labels	Rankings	Criterion priorities	Assessment	Content	Effort	Enjoyment	Faculty	Interaction	Structure	Usefulness
A1	3	0.10780717	0.08041060	0.04486510	0.07660304	0.07020630	0.24214347	0.07088712	0.06318701	0.30648066
A2	2	0.12317375	0.14738911	0.17332193	0.10660947	0.16427961	0.13802755	0.28743492	0.20895735	0.04484990
A3	1	0.14937927	0.20189425	0.23792359	0.14717195	0.18429596	0.06114298	0.15602703	0.11956994	0.05324379
A4	5	0.10102999	0.12193964	0.18535673	0.09115032	0.20614247	0.08324288	0.13764311	0.12511872	0.18203878
A5	6	0.08912817	0.12531078	0.11357968	0.07752716	0.12260709	0.05432878	0.10733931	0.10203971	0.06857601
A6	8	0.07043324	0.06178207	0.05347288	0.12691945	0.07815893	0.05957748	0.07648407	0.05784910	0.09802823
A7	9	0.06145910	0.02596659	0.02854271	0.08162745	0.03684313	0.07320527	0.04614353	0.16821623	0.06537621
A8	7	0.07999366	0.03386392	0.06500947	0.08919895	0.05395722	0.06321575	0.05325129	0.05858617	0.05830051
A9	4	0.10512918	0.03560174	0.06562576	0.08965879	0.05033126	0.19985944	0.04001535	0.04555374	0.11894135

### 5.2 Course ranking

The scores for each criterion in the case of Art, Design, and Humanities courses are presented in Table 4. Eight sets of pairwise comparison matrices were utilized to evaluate and rank the quality of courses within Art, Design, and Humanities, as listed in Tables 5 and 6, and Tables 12, 13, 14, 15, 16, 17 and 18 in the “Appendix”. Let the  $n$ -by- $n$  pairwise comparison matrix be denoted as  $A = (a_{ij})$ , where  $i, j = 1, 2, \dots, n$ . The entry  $a_{ij}$  is determined by  $w_i/w_j$ , where  $w_i$  and  $w_j$  represent the weights of alternatives  $i$  and  $j$  with respect to certain criteria. For the purpose of comparison, the nine courses are enumerated on the left-hand side and at the top, and an evaluation is conducted to assess the degree of dominance of courses on the left over those at the top. For instance, in Table 5, when we compare course A1 on the left with course A2 at the top with respect to the “assessment” criterion, we cross-reference Table 4, where the “assessment” ratings for A1 and A2 are 93.058 and 170.572. Consequently, we record the value 0.546 (93.058/170.572) in the cell located at the intersection of the first row and second column in Table 5, while the value 1.833 (170.572/93.058) is automatically populated in the cell at the intersection of the second row and first column in Table 5. In this scenario, it is observed that the quality of course A1, specifically regarding the “assessment” criterion on the left, does not surpass that of course A2 at the top.

To gain the priorities for each table, this study calculated the principal right eigenvector for each matrix and normalized its values by dividing each entry by the total sum. This allows ranking each course according to its priorities for each criterion. The last column in Table 5 indicates the ranking position of the nine courses in the case of the criterion “assessment”. The ranking results were determined by the priorities of the criterion of “assessment” presented in the penultimate column of Table 5. The results showed that A3 is ranked at the top regarding “assessment”, followed by A2, and A5, and A7 is ranked as the lowest.

In Table 6, this study determined the priorities by computing the sum score weighted accordingly. For instance, the priorities of A1 are calculated as follows: 0.09236552 (the weight of “assessment”)  $\times$  0.0804106 + 0.0995969 (the weight of “content”)  $\times$  0.0448651 + 0.12108072 (the weight of “effort”)  $\times$  0.07660304 + 0.02847555 (the weight of “enjoyment”)  $\times$  0.0702063 + 0.26967681 (the weight of “faculty”)  $\times$  0.24214347 + 0.05056552 (the weight of “interaction”)  $\times$  0.07088712 + 0.03175832 (the weight of “structure”)  $\times$  0.06318701 + 0.30648066 (the weight of “usefulness”)  $\times$  0.0448499 = 0.10780717. Our method ranked alternatives within Art, Design, and Humanities course selection, as presented in column 2 of Table 6. The ranking results were determined by the priorities in column 3 of

**Table 7** Score for each criterion in Computer Science, Engineering, and Programming course selection

	Enjoyment	Structure	Faculty	Assessment	Content	Interaction	Usefulness	Effort	Technicality
B1	72.525	36.709	76.349	105.882	50.420	21.799	43.911	92.937	34.945
B2	76.157	41.779	116.268	105.142	46.802	31.749	53.869	68.923	31.818
B3	264.714	145.642	301.376	376.166	156.838	108.667	174.095	278.110	135.768
B4	76.553	43.732	138.705	100.787	51.603	28.740	46.650	67.044	30.696
B5	78.489	44.526	110.906	91.837	41.024	30.594	54.437	100.028	35.906
B6	65.039	33.160	67.655	72.648	34.664	20.082	38.365	68.870	26.958
B7	117.328	65.143	240.424	189.968	77.221	51.187	87.350	131.530	52.551
B8	195.385	100.997	249.214	182.237	88.089	55.774	121.712	169.002	63.775
B9	77.716	44.248	97.315	107.034	52.705	27.123	52.853	71.649	35.377

**Table 8** Pairwise comparison matrix for alternatives regarding enjoyment in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	0.952	0.274	0.947	0.924	1.115	0.618	0.371	0.933	0.07083156	8
B2	1.050	1	0.288	0.995	0.970	1.171	0.649	0.390	0.980	0.07437841	7
B3	3.650	3.476	1	3.458	3.373	4.070	2.256	1.355	3.406	0.25853377	1
B4	1.056	1.005	0.289	1	0.975	1.177	0.652	0.392	0.985	0.07476602	6
B5	1.082	1.031	0.297	1.025	1	1.207	0.669	0.402	1.010	0.07665607	4
B6	0.897	0.854	0.246	0.850	0.829	1	0.554	0.333	0.837	0.06352012	9
B7	1.618	1.541	0.443	1.533	1.495	1.804	1	0.600	1.510	0.11458888	3
B8	2.694	2.566	0.738	2.552	2.489	3.004	1.665	1	2.514	0.19082333	2
B9	1.072	1.020	0.294	1.015	0.990	1.195	0.662	0.398	1	0.07590184	5

Table 6, showing that A3 is ranked at the top, followed by A2, and A1, and A7 is ranked as the lowest.

Analogous computations are conducted to determine the rankings for the alternatives in the selection of Computer Science, Engineering, and Programming courses, as displayed in Table 7. A total of 9 pairwise comparison matrices are formulated using data from Table 8 and Tables 19, 20, 21, 22, 23, 24, 25 and 26 in the “Appendix”, aiming to evaluate the Computer Science, Engineering, and Programming courses based on the criteria of “enjoyment”, “structure”, “faculty”, “assessment”, “content”, “interaction”, “usefulness”, “effort”, and “technicality”. For example, the last column in Table 8 indicates the ranking position of the nine courses in the case of the criterion of “enjoyment”. The ranking results were determined by the priorities of the criterion “enjoyment” presented in the penultimate column of Table 8. The results showed that B3 is ranked at the top regarding “enjoyment”, followed by B8, and B7, and B6 is ranked as the lowest.

The rankings for alternatives are presented in column 2 in Table 9 for Computer Science, Engineering, and Programming course selection. The ranking results were determined by the priorities in column 3 of Table 9, showing that B3 is ranked at the top, followed by B8, and B7, and B6 is ranked as the lowest.

## 6 Discussion

### 6.1 RQ1: What factors will be included in the multi-criteria decision-making framework? What are the similarities and differences between the selection of knowledge- and skill-seeking courses?

This study revealed notable distinctions between subjects. The decision hierarchies presented in Figs. 4 and 5 offer valuable insights into the factors essential for multi-criteria decision-making in course selection. For both Art, Design, and Humanities (D1) and Computer Science, Engineering, and Programming (D2) courses, eight criteria: “assessment”, “content”, “effort”, “usefulness”, “enjoyment”, “faculty”, “interaction”, and “structure” were considered. These criteria are widely acknowledged for their impact on MOOC learners' enrollment, satisfaction, and reuse intention [32, 36, 37].

The sole distinction lies in the inclusion of “technicality” for Computer Science, Engineering, and Programming courses. The “technicality” refers to the non-monetary cost related to the use of technology and plays a crucial role in users' technology adoption [38]. It consists of two sub-criteria: “complexity” and “flexibility”. Complexity pertains

**Table 9** Numerical values for ratings of the alternatives in Computer Science, Engineering, and Programming course selection

Labels	Rankings	Criterion priorities	Employment	Structure	Faculty	Assessment	Content	Interaction	Usefulness	Effort	Technicality
B1	8	0.07137678	0.07083156	0.06603153	0.05460495	0.07950892	0.08412255	0.05801926	0.06522387	0.08867243	0.05169002
B2	6	0.07677539	0.07437841	0.07515078	0.08315478	0.07895288	0.07808652	0.08450420	0.08001422	0.06576027	0.07105562
B3	1	0.25903943	0.25853377	0.26197673	0.21554366	0.28247013	0.26167295	0.28922762	0.25859129	0.26534916	0.30319182
B4	5	0.07850634	0.07476602	0.07866349	0.09920132	0.07568302	0.08609547	0.07649307	0.06929152	0.06396742	0.06854996
B5	4	0.07906708	0.07665607	0.08009194	0.07931976	0.06896232	0.06844490	0.08142826	0.08085850	0.09543809	0.08018378
B6	9	0.05744728	0.06352012	0.05964629	0.04838699	0.05455261	0.05783515	0.05345036	0.05698581	0.06570977	0.06020207
B7	3	0.13570307	0.11458888	0.11717687	0.17195107	0.14265084	0.12883743	0.13623914	0.12974468	0.12549462	0.11735581
B8	2	0.16635332	0.19082333	0.18167006	0.17823778	0.13684519	0.14696976	0.14844739	0.18078532	0.16124730	0.14242091
B9	7	0.07573131	0.07590184	0.07959232	0.06959969	0.08037410	0.08793528	0.07219070	0.07850481	0.06836092	0.07900020

to the ease of use of technology, wherein a user-friendly system requiring minimal physical and mental effort leads to a smoother and more satisfying learning experience. Flexibility involves the adaptability of MOOCs regarding time, space, and learning pace, enabling learners to personalize their learning experience, which enhances their satisfaction [39].

Regarding the eight shared criteria between D1 and D2 courses, inconsistencies were observed in the sub-criteria. For “assessment”, D1 only includes “assignment” as a sub-criterion, while D2 adds “autograder”. This finding supports the research by Ramesh et al. [40], highlighting the significance of auto-gradable hands-on programming assignments in scalable programming education. The “autograder”, providing automatic feedback on programming exercises, enhances learners' satisfaction and understanding of code [41].

Another disparity lies in the inclusion of “explanation” as a sub-criterion under the “content” criterion for D2, which is not present in D1. Learners in Computer Science, Engineering, and Programming courses value the clear explanation of complex concepts, making it essential for their continued learning [42]. The ability to explain such abstract scientific concepts in an accessible manner is crucial, emphasizing the importance of problem-based learning and explicit explanations to enhance participation in online courses [43].

Furthermore, an illustrative case is the consideration of instructor humor as a sub-criterion within the “faculty” category for D2, while it is absent in D1. According to Watson et al. [44], this aspect of instructors, humor, is thought to contribute to heightening the emotional aspect of social interaction, referring to the extent to which users using computer-assisted communication systems experience an emotional connection with one another [45]. Extensive literature elucidates the significant influence of both instructor presence and social presence on learners' satisfaction with online learning (e.g., [46, 47]). In the realm of Computer Science, Engineering, and Programming courses, where abstract and enigmatic concepts often surface, instructors' humor plays a particularly crucial role in inspiring learners to actively engage with and grasp these notions.

**6.2 RQ2: What are the most influential factors in course selection? What are the similarities and differences between the selection of knowledge- and skill-seeking courses?**

Among the criteria with a weight of more than 10%, “effort” and “faculty” appeared in both D1 and D2. When considering distinctions, the evaluations of the criteria

“assessment”, “enjoyment”, and “usefulness” differ between D1 and D2.

### 6.2.1 Effort

Regarding Art, Design, and Humanities courses, the “effort” criterion contributes 12.11% to the overall evaluation process, while for the selection of Computer Science, Engineering, and Programming courses, it carries a weight of 14.03%.

The effort is inherently intertwined with the technological aspects mentioned previously. As MOOCs are designed with multifaceted functionalities to meet users’ expectations, their adaptation requires certain extents of mental and physical exertion, for example, time, energy, financial commitment, engagement, and utilization of resources. Through our topic modeling analysis, we identified perceived difficulty and perceived workload as indicators of perceived effort.

Perceived workload pertains to students’ perception of the workload and time commitment required for completing course tasks [32]. Perceived difficulty refers to a student’s perception of how easy or challenging it is to comprehend the course content [48]. Before enrolling, learners often evaluate the course description, syllabus, and any available information about the workload and course difficulty. When learners perceive the workload and course difficulty as high, they are likely to perceive the effort required to complete tasks as substantial. Consequently, if the perceived workload and course difficulty align with their expectations, available time, and resources, learners are more inclined to enroll. However, if learners anticipate a high workload and course difficulty that surpasses their capacity or conflicts with their goals, it can result in increased stress, anxiety, and reduced satisfaction, and may impact learners’ decisions to enroll in a MOOC. Our findings align with the observations in typical in-person classroom settings. Howell and Buck [49] reported a significant association between workload and student satisfaction and engagement in face-to-face learning. Nonetheless, the correlation between perceived workload and perceived effort may vary depending on individual factors, such as prior knowledge and learning preferences.

### 6.2.2 Faculty

Regarding the “faculty” criterion, it contributes 26.97% and 19.63% to the evaluation of Art, Design, and Humanities, and Computer Science, Engineering, and Programming courses, respectively. Our topic modeling analysis results show that within the context of MOOCs, the “faculty” criterion consists of six sub-criteria: instructor experience, faculty team, instructor passion, instructor humor, instructor presentation, and teaching style. These factors have been

extensively emphasized in the literature as crucial elements in MOOCs (e.g., [6, 50, 51]).

The significance of “faculty” as a factor influencing MOOC learners’ satisfaction can be attributed to several reasons. Firstly, experienced instructors possess vast knowledge and expertise, effectively guiding students, addressing queries, and providing real-world examples. This enriches the learning experience and fosters learner satisfaction. Secondly, a cohesive and supportive faculty team creates a positive environment. MOOC platforms offer interactive tools that assist the faculty team, including instructors and teaching assistants, in problem-solving and facilitating learning inquiry [52], thereby enhancing personalized attention to students. Thirdly, passionate instructors, go beyond the curriculum, and inspire and motivate students. Their enthusiasm encourages exploration and deeper understanding, resulting in an engaging and enjoyable learning experience and heightened satisfaction [53]. Fourthly, the appropriate use of humor by instructors creates a relaxed atmosphere, alleviating tension, fostering rapport, and establishing a positive emotional connection with learners. These positive emotions contribute to a more satisfying learning experience. Fifthly, effective presentation skills, such as clear communication, well-organized lectures, and the skillful use of visual aids, significantly impact learner satisfaction. Instructors who deliver information coherently and engagingly facilitate the comprehension of complex concepts and maintain student interest. Additionally, accommodating diverse learning styles through adaptable teaching methods creates an inclusive and effective learning environment [54]. By catering to individual needs and employing various instructional approaches, instructors enhance learner satisfaction by promoting understanding and practical application of course material.

Overall, the “faculty” criterion encompasses various factors that directly influence instructional quality and the overall learning experience. Fulfilling these criteria enhances learner engagement, motivation, and satisfaction in MOOCs.

### 6.2.3 Assessment

In terms of the criterion “assessment”, its weight is 16.19% in D2, while it constitutes only 9.24% of the evaluation process in D1. In the context of MOOCs, the significance of assessment is widely acknowledged. For example, Bali [55] proposed that MOOC assessment provides learners with opportunities to evaluate their acquired knowledge and apply their learned skills. Likewise, Hew [43] emphasized the significance of interactive learning and applying knowledge to improve engagement. The correlation between motivation and evaluation is supported by Zimmerman [56],

who discovered that individuals with particular learning objectives tend to assess their performance. Scholars have also observed that providing self-assessment tools to learners aids in developing critical thinking abilities [57].

Especially in Computer Science, Engineering, and Programming MOOCs, learners often discuss assessment-related matters. This aligns with the considerable interest among researchers in exploring effective assessment strategies to enhance computer science and programming education [58]. Although auto-graders in MOOCs have improved fairness and efficiency, challenges remain, such as potentially inaccurate auto-grading and a lack of personalization. To optimize programming education, it is suggested to support inline-anchored discussions addressing auto-grader complaints and integrate program visualization tools that demonstrate how auto-graders work into forums. Additionally, the assessment features characteristic of MOOCs, such as assignments and quizzes, still fall short of meeting the requirements of effective computer science and programming instruction. To foster proficient programming skills, learners must receive timely and accurate feedback on their solutions and answers after completing practical programming exercises and real-world programming tasks.

#### 6.2.4 Enjoyment

Concerning the criterion “enjoyment” its weight is 15.81% in D2, while constituting only 2.85% of the evaluation process in D1. Cognitive Evaluation Theory suggests that individuals’ actions are shaped by internal components, emphasizing the joy and contentment derived from engaging in behavior [59]. Enjoyment, referring to “the extent to which the activity of using a product is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated (p. 1113)” [60], is considered an inherent and emotional concept that impacts teachers’ acceptance. This concept comprises two sub-components: pleasure and satisfaction, gauging teachers’ internal drive across various levels of necessities.

Pleasure represents the initial level of positive mental states learners experience in learning [33]. Although MOOCs are not primarily for pleasure-driven intentions, learners may find it interesting, thus motivating continued adoption. Satisfaction is an emotional reaction resulting from achieving a certain objective after usage [61]. When learners are satisfied with the learning outcomes from a MOOC, they have higher intentions for continued use. Howarth et al. [62] found a positive correlation between

current learners’ satisfaction and technology adoption by new learners.

In Computer Science, Engineering, and Programming MOOCs, enjoyment plays a pivotal role for two reasons. First, these fields demand high levels of cognitive effort and problem-solving skills. When learners find the activities and learning experiences intrinsically enjoyable, it enhances their motivation to engage with the MOOCs. Enjoyment sustains their interest, curiosity, and enthusiasm, fostering higher levels of engagement and adoption. Second, Computer Science, Engineering, and Programming involve complex concepts, abstract thinking, and technical skills. The learning process can be demanding and require significant effort. Enjoyment serves as a positive emotional response that helps alleviate the perceived difficulty and challenges associated with these subjects. When learners derive joy and pleasure from the learning activities, they are more likely to persist through challenges, remain motivated, and continue adopting MOOCs.

#### 6.2.5 Usefulness

Regarding usefulness, it constitutes 30.65% of the evaluation process in D1, while accounting for only 9.98% in D2. Utility, characterized as an external and mental concept, refers to the total worth a user attributes to employing a new technology [38]. Through our topic modeling analysis results, in the context of MOOCs, the utility has three sub-criteria: knowledge enhancement, job preparedness, and beginner friendliness. Among these, job preparedness is exclusively present in D1, making up 16.81% of the evaluation process. This somewhat supports [63], who identified job promotion needs and curiosity as primary factors influencing learners’ enrollment in a MOOC. However, motives for MOOC enrollment can vary across different course types. For example, according to a survey by Bayeck [64], 74.6% and 11.9% of respondents enrolled in humanities courses (poetry and music, respectively) due to curiosity and seeking job performance improvement. For MOOCs related to science, health science, and math, only 39% of enrollments were motivated by skill improvement to perform better in jobs.

For learners taking Art, Design, and Humanities MOOCs, potential explanations for their frequent concerns about job preparedness are as follows. Firstly, learners in Art, Design, and Humanities disciplines often prioritize acquiring practical skills that directly apply to their work. They seek assurance that the courses they undertake will equip them with relevant and applicable skills sought after in the job market. It is essential to connect theoretical understanding with real-world implementation, ensuring

that the courses they invest time and effort in will have a tangible impact on their professional growth. Secondly, the Art, Design, and Humanities industries are diverse and encompass various sectors, each with its unique requirements and expectations. Learners are concerned about whether the MOOCs they choose align with their career interests and goals. They want to ensure that the course content, projects, and assignments offered reflect current industry practices. The relevance of the course content to their chosen field is crucial for learners to feel confident in job preparedness.

### 6.3 RQ3: What are the ratings of courses and how do they rank according to sub-criteria and criteria?

The overall rankings of course alternatives in selecting Art, Design, and Humanities as well as Computer Science, Engineering, and Programming courses, as presented in Tables 6 and 9, do not precisely match the overall course ratings from the Class Central website (Tables 2 and 3). There are similarities and differences between them. For instance, according to the reviews' overall ratings about Art, Design, and Humanities courses in column 3 of Table 2, A2, A5, and A8 are rated as 4.7, which is the lowest among the nine courses. However, our proposed model, as shown in column 2 of Table 6, ranks A5 and A8 in the seventh and eighth positions, respectively, while it places A2 in the second position. Similarly, for the Computer Science, Engineering, and Programming courses, based on the reviews' overall ratings in column 3 of Table 3, B2 and B8 receive the highest rating of 4.9 among the nine courses. Nevertheless, our proposed model, displayed in column 2 of Table 9, ranks B8 in the second position and B2 in the sixth position.

The inconsistencies between our analysis results and the reviewers' overall opinions about the courses may arise from the fact that our approach goes beyond merely considering rough scores given by individual learners. Instead, we employ a more detailed and fine-grained approach by also considering the weights and significance attached to each criterion, determined through group decision-making analysis. This allows us to gain insights into a single course's rankings based on different criteria and even sub-criteria, which are not available on the Class Central website. Results presented in Table 5 and Tables 12, 13, 14, 15, 16, 17 and 18 in the "Appendix" for Art, Design, and Humanities course selection, and results presented in Table 8 and Tables 19, 20, 21, 22, 23, 24, 25 and 26 in the "Appendix" for Computer Science, Engineering, and Programming course selection, provide a more comprehensive understanding of how courses perform according to various criteria and

sub-criteria. In contrast, the same overall course rating presented in Tables 2 and 3 treats all courses with the same overall rating as equal.

For example, in the case of Art, Design, and Humanities, we ranked "English for the Workplace" (A3) as the best course. By checking Table 5 and Tables 12, 13, 14, 15, 16, 17 and 18 in the "Appendix", we find that A3 is also ranked the first among all alternatives according to four criteria: "assessment", "content", "effort", and "usefulness". The weights of these four criteria, from Table 6, are approximately 9.24%, 9.96%, 12.10%, and 30.64%, respectively, collectively constituting about 61.95% of the evaluation process. Therefore, we ranked "English for the Workplace" in the first place. For the other four criteria: "enjoyment", "faculty", "interaction", and "structure", the courses ranked the first among all alternatives were "English in Early Childhood: Language Learning and Development" (A4), "A Life of Happiness and Fulfillment" (A1), and "Academic Writing" (A2), respectively. Although both A1 and A4 are rated 4.7 on the Class Central website, our analysis shows that they perform differently based on different evaluation criteria. By having this knowledge, learners can select courses with better performance in the criteria they value the most. Therefore, implementing the AHP enables us to distinguish between courses that have the same overall course rating obtained from the Class Central website.

### 6.4 Implications and suggestions

The results of this investigation illuminate the factors that influence the selection of MOOC courses and offer valuable perspectives for instructors and researchers to optimize course design, instructor involvement, and learner satisfaction in MOOCs.

1. Course design and presentation
 

Instructors of MOOCs should create courses that go beyond being informative and are also captivating, flexible, and well-structured. Moreover, for skill-oriented courses, integrating technical aspects like user-friendliness and adaptability is crucial to enhance learners' experiences.
2. Instructor qualities and engagement
 

Instructors should cultivate a positive learning environment and motivate learners through their expertise, experience, and enthusiasm. Incorporating humor into instruction can enhance social presence and emotional connection. They should deliver clear, well-organized lectures, effectively utilize visual aids, and embrace adaptable teaching methods catering to diverse learning styles.



### 3. Assessment strategies

Instructors need to carefully consider assessment methods, particularly in skill-oriented courses where auto-graders play a vital role in providing timely feedback and increasing satisfaction. However, they must address potential issues of inaccurate grading and lack of personalization. Introducing program visualization tools and fostering discussions can be beneficial. Assessments should promote active learning and self-evaluation to foster critical thinking skills.

### 4. Promoting enjoyment

Instructors should prioritize creating enjoyable learning experiences, especially in skill-oriented MOOCs like Computer Science, Engineering, and Programming, which can be challenging. Enjoyment positively impacts learners' emotions, reduces perceived difficulty, and encourages engagement. Interactive and engaging activities enhance motivation and satisfaction.

### 5. Addressing job preparedness

In knowledge-oriented MOOCs like Art, Design, and Humanities, addressing learners' concerns about job preparedness is crucial. Instructors should ensure courses offer practical and applicable skills aligned with learners' career interests. Demonstrating the relevance of course content to specific industries and job requirements boosts learners' confidence and satisfaction.

### 6. Enhancing learner engagement

MOOC platforms should provide personalized course recommendations based on individual preferences and learning styles to enhance engagement. Tailoring courses to meet learners' valued criteria improves satisfaction and retention. Additionally, allowing learners to evaluate courses based on different criteria facilitates informed decision-making.

### 7. Leveraging multi-criteria decision-making

Researchers and MOOC platforms can employ the multi-criteria decision-making methodologies introduced in this study for comprehensive course evaluations. By considering overall course ratings and individual criteria weights, learners can make informed decisions aligned with their preferences. Researchers can refine the model by adding factors to improve evaluations.

By implementing these insights and recommendations, educators and online learning platforms can enhance the effectiveness and appeal of MOOC experiences, thereby promoting the advancement and enhancement of online education.

## 6.5 Reflections on research methodologies

This study used courses with an overall rating score of 4.7 or higher based on observing user behaviors on real-world online platforms. Typically, users gravitate towards courses with higher overall ratings when making selections. The challenge lies in the limited information displayed—often only the overall course ratings are visible on the MOOC website. Despite the presence of numerous textual reviews from previous learners, manually evaluating each comment becomes a time-consuming task. Additionally, learners have diverse priorities, some favoring assessment aspects while others value interaction. This diversity contributes to the difficulty users face in objectively evaluating and selecting courses based on the available information. Our method aims to encourage users to delve deeper into the evaluation process for the initially selected highly rated courses. By providing more granular evaluation criteria and insights into various performance dimensions, learners can identify courses that align closely with their preferences. Thus, we specifically focused on comparative analysis among courses with high overall rating scores across different dimensions, mirroring real-world user decision-making scenarios. This is more practical and meaningful, and aligns more closely with user needs during the course selection process, particularly when dealing with popular courses exhibiting minimal differences in overall ratings. However, it would be intriguing in future endeavors to include a larger range of overall rating scores of courses for comparison.

Regarding the predetermined labels used for guiding the labeling process of topics, while we did not explicitly create standalone labels for “videos” and “instructors” that are frequently used in previous studies (e.g., [2, 32]), we have covered various dimensions pertinent to instructors and videos within the “technicality” and “faculty” labels. Specifically, the structuring of the “technicality” label, with its comprehensive sub-labels, was determined to encompass and evaluate the essential elements pertaining to technological factors including video content in MOOCs. By analyzing the technical intricacies (complexity) and adaptive features (flexibility) of MOOC videos, we ensured a holistic evaluation that inherently included the pivotal role and influence of video content on the overall quality of MOOC courses. Furthermore, the structuring of the “faculty” label, with its detailed sub-labels, was utilized to thoroughly encompass the multiple dimensions associated with instructors in MOOC courses. Our approach aimed at holistically evaluating the influence of instructors by dissecting various aspects that collectively contribute to the overall perception of faculty quality within the MOOC setting. By employing the “faculty” label and its sub-labels, we aimed to cover a broad spectrum of factors that directly

relate to instructors' effectiveness, teaching styles, and collaborative dynamics within the teaching team, thereby providing a comprehensive evaluation. Nevertheless, future work can consider proposing more fine-grained methods, for example, the Delphi method by involving multiple rounds of questionnaire surveys to expert groups, to determine predetermined labels.

Regarding the manual analysis of the algorithm-identified topics, this is a prevalent method in topic modeling studies (e.g., [16, 65]). This study rigorously adhered to a widely adopted procedure from previous studies (e.g., [34, 35]). Specifically, this study's interpretation of topics was guided by sociotechnical system theory, Moore's transactional distance theory, and MOOC instructional design practices. To enhance precision, our methodology entailed multiple rounds of in-depth discussions among two experienced MOOC education experts until a consensus was reached. While involving two annotators is generally considered to be able to ensure a higher level of accuracy and reliability in the labeling process in previous topic modeling studies (e.g., [35, 66]), we acknowledge the potential for increased rigor by involving more annotators. Moreover, future work might consider integrating STM outcomes with automated methods for topic naming, thereby improving the interpretative process's accuracy and efficiency.

Regarding the determination of models based on exclusivity and coherence scores across a range of 5 to 50 topics, this approach might appear simplistic in finding the most suitable number of topics. However, exclusivity and coherence measures are widely considered effective methods in previous topic modeling studies (e.g., [65, 67]) to define the ideal number of topics. These measures serve as heuristics by providing indicators or scoring systems that help evaluate the quality or coherence of topics generated by models. They aid in the selection of an optimal number of topics by assessing the semantic coherence within topics and the distinctiveness between them [68]. Nevertheless, future work can consider involving and combining more advanced heuristic-driven methodologies such as perplexity metrics and potentially exploring Bayesian approaches to achieve a more refined and accurate identification of the ideal number of topics.

Regarding the computational complexity, the optimization of topic models entails the refinement of parameters and the selection of an appropriate model configuration. Specifically, a substantial amount of computing resources is dedicated to determining the most suitable models based on exclusivity and coherence

scores across a range of topics (from 5 to 50) within the proposed method, requiring hours for processing databases D1 and D2. The computation of exclusivity and coherence scores for various topic numbers involves running multiple iterations of the topic modeling algorithm for each specified number of topics. This iterative process is computationally intensive, in comparison to other steps in the proposed method, for example, the automatic ranking of a sample of courses in various criteria, which can be achieved within seconds. Nevertheless, the overall computational time for the proposed method is much shorter than using manual coding. However, as dataset sizes expand, encompassing larger vocabularies and broader ranges of topics, the time and resources for model training and inference will also increase, potentially leading to scalability concerns. Thus, future endeavors may contemplate implementing optimization techniques within scoring algorithms or adopting more computationally efficient approaches for computing exclusivity and coherence scores to alleviate the computational burden inherent in the overall modeling process.

Overall, the proposed methodology in this study enables effective and efficient analysis of extensive textual data, a stark contrast to manual analysis methods. However, we recognize the limitation compared to a "true close reading" of texts. Yet, automated text analysis offers tremendous potential for educational researchers to grapple with vast volumes of data, facilitating the exploration of MOOC learner-generated texts.

## 6.6 Limitations and future work

This study has limitations. Specifically, the review data was exclusively obtained from the Class Central, a well-known MOOC aggregator website. While the Class Central is esteemed, incorporating data from diverse MOOC platforms such as EdX, Coursera, self-developed online learning platforms, and even open-ended MOOC surveys from learners and instructors would offer additional validation and complement the study's findings. The present study solely concentrated on Art, Design, and Humanities as well as Computer Science, Engineering, and Programming MOOCs as representatives of knowledge- and skill-seeking courses. To enhance the applicability and generalization of the findings, it is imperative to include MOOCs from a broader spectrum of courses, thereby augmenting the model's effectiveness. Since courses from different disciplines may display unique characteristics.

## 7 Conclusion

This study, introduced a novel approach to automatically evaluate MOOCs using a multi-criteria decision-making model, combining text mining and AHP. The research analyzed reviews from both knowledge-seeking MOOCs, such as Art, Design, and Humanities, and skill-seeking MOOCs, such as Computer Science, Engineering, and Programming courses. Common criteria prioritized by both types of learners include “assessment”, “content”, “effort”, “usefulness”, “enjoyment”, “faculty”, “interaction”, and “structure”. Additionally, skill-seeking learners also emphasized technicality, encompassing ease of use and flexibility. For both knowledge-seeking and skill-seeking learners, effort and faculty were deemed crucial. However, skill-seeking learners placed more importance on “assessment” and “enjoyment”, while knowledge-seeking learners valued “usefulness” higher. The research demonstrated the effectiveness of employing online learner reviews and topic modeling for automated MOOC evaluation.

Reflecting on the research design of this study, several issues should be noted: (1) use of data related only to Art, Design, Humanities, Computer Science, Engineering, and Programming courses with an overall rating score of 4.7 or higher from Class Central, (2) determination of predetermined labels based on a review of literature, (3) manual analysis of the algorithm-identified topics, and (4) determination of models based on exclusivity and coherence scores across a range of 5 to 50 topics. Future work may pay attention to the following issues. First, including a larger range of overall rating scores of courses from different disciplines and diverse

MOOC platforms. Second, proposing more fine-grained methods such as the Delphi method to determine predetermined labels. Third, integrating STM outcomes with automated methods for topic naming. Fourth, combining more advanced heuristic-driven methodologies such as perplexity metrics and potentially exploring Bayesian approaches to determine the ideal number of topics. Additionally, it is also vital to acknowledge that latent topics within course reviews may evolve over time. Thus, future research could focus on developing tools capable of providing real-time analyses of the topics relevant to MOOC learners.

In sum, the proposed automated text analysis methodology in this study offers tremendous potential for educational researchers’ effective and efficient analysis of vast volumes of textual data to shed light on course evaluation and selection. The findings of this study have significant implications for optimizing course design, enhancing instructor engagement, and improving learner satisfaction. MOOC platforms can leverage this model to provide personalized course recommendations, ultimately contributing to the advancement of online education and enhancing the MOOC learning experience for learners.

## Appendix

See Tables 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 and 26.

**Table 10** Extracted valid topics and their adjusted weights and discriminant terms for D1 (Art, Design, and Humanities)

Labels	Adjusted weights (%)	Discriminant terms
Interaction	5.06	Forum, prepared, deliver, interaction, responsive, presentation, clearly, lecture, support
Perceived difficulty	5.88	Understand, easy, medium, concept, chapter, simple, ease, comprehend, difficulty, step
Perceived workload	6.23	Work, hard, half, day, problem, busy, submit, week, schedule, date
Beginner friendliness	8.15	Beginner, basic, basis, slow, speak, production, pace, begin, start, introductory
Instructor experience	13.47	Lesson, teacher, teaching, plan, strategic, idea, method, experience, experienced, technique
Job preparedness	16.81	Job, interview, opportunity, skill, Facebook, prepare, workplace, useful, Futurelearn, join
Knowledge enhancement	5.69	Knowledge, sustainability, gain, expand, industry, broaden, business, industrial, expertise, professionally
Assignment	9.24	Exercise, exam, complete, certificate, pay, test, correct, upgrade, fee, spend
Problem-solving and use of cases/examples	9.96	Explain, question, ask, example, answer, comprehensive, case, clear, reference, tip
Faculty team	6.56	Faculty, scholar, research, coordinator, researcher, team, Ugc, presentation, college, institute
Course description	3.18	University, critical, history, Sheffield, middle, Tsinghua, degree, French, aid, existentialism
Instructor passion	6.94	Happiness, happy, prof, fulfillment, professor, joy, energy, entertaining, fulfilling, passion
Pleasure	2.85	Interesting, enjoy, information, video, topic, interest, short, time, online, useful

**Table 11** Extracted valid topics and their adjusted weights and discriminant terms for D2 (Computer Science, Engineering, and Programming)

Labels	Adjusted weights (%)	Discriminant terms
Pleasure	4.79	Engaging, fantastic, enjoyable, entertain, top, pleasure, entertaining
Course introduction	7.92	Everybody, first, continue, one, series, university, Michigan, Coursera, specialization, specialisation
Instructor presentation	5.71	Teacher, drchuck, didactic, interesting, simpler, interest, brilliant, extraordinary, funny, profchuck
Teaching style	6.73	Love, awesome, teaching, amaze, bore, amazing, granulation, teach, style, join, cool, ease, lovely, stuff, fabulous
Autograder	3.84	Command, report, editor, autograder, write, copy, code, mistake, debut, correct, output, prompt, execute
Satisfaction	11.02	Great, recommend, highly, strongly, interested, recommended, cheer, definitely, wanting
Problem solving	2.68	Wrong, professional, error, mark, ask, address, question, answer, audit, submit, capability, breakpoint, aid
Interaction	4.25	Help, platform, revise, staff, team, discussion, forum, community, assist, solve
Beginner friendliness	6.57	Beginner, useful, helpful, level, perfectly, experienced, syllabus, smoothly, appropriate, suit
Perceived difficulty	9.56	Program, basic, start, know, already, nothing, background, conditional, fundamental, familiar, afraid, loop, worry
Instructor humor	3.11	Humor, fun, sense, achievement, laugh, interactive, pleasant, joyful, exciting, instructive, tremendous, joy, atmosphere
Flexibility	2.66	Section, put, formal, hit, habit, view, remind, audience, somewhere, remote, waste, drop, record
Knowledge enhancement	3.41	Knowledge, gain, acquire, refresh, prior, zero, experience, expertise, little, gap, fill
Explanation	4.34	Clearly, detail, explain, complete, every, explained, layman, able, breeze, patiently, comprehensible, neatly, illustrate
Perceived workload	4.47	Week, exam, confuse, speed, spend, hour, weekend, invest, rush, deadline, longer, spent, spare, duration, trouble, hurry, devote
Assessment	12.36	Assignment, easy, exercise, practice, concise, quire, quit, test, tough, difficult, reinforce, tricky, digest, assessment, quiz
Instructor passion	4.07	Prof, humorous, engage, professor, passionate, extremely, enthusiastic, patient, charismatic, passion, attractive, explicit, genuinely, guidance
Complexity	2.51	Think, something, point, need, try, find, still, yet, complex, progress, felt, hard

**Table 12** Pairwise comparison matrix for the alternatives regarding content criterion in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.259	0.189	0.242	0.395	0.839	1.572	0.690	0.684	0.04486510	8
A2	3.863	1	0.728	0.935	1.526	3.241	6.072	2.666	2.641	0.17332193	3
A3	5.303	1.373	1	1.284	2.095	4.449	8.336	3.660	3.625	0.23792359	1
A4	4.131	1.069	0.779	1	1.632	3.466	6.494	2.851	2.824	0.18535673	2
A5	2.532	0.655	0.477	0.613	1	2.124	3.979	1.747	1.731	0.11357968	4
A6	1.192	0.309	0.225	0.288	0.471	1	1.873	0.823	0.815	0.05347288	7
A7	0.636	0.165	0.120	0.154	0.251	0.534	1	0.439	0.435	0.02854271	9
A8	1.449	0.375	0.273	0.351	0.572	1.216	2.278	1	0.991	0.06500947	6
A9	1.463	0.379	0.276	0.354	0.578	1.227	2.299	1.009	1	0.06562576	5

**Table 13** Pairwise comparison matrix for the alternatives regarding effort in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.719	0.521	0.840	0.988	0.604	0.938	0.859	0.854	0.07660304	9
A2	1.392	1	0.724	1.170	1.375	0.840	1.306	1.195	1.189	0.10660947	3
A3	1.921	1.380	1	1.615	1.898	1.160	1.803	1.650	1.641	0.14717195	1
A4	1.190	0.855	0.619	1	1.176	0.718	1.117	1.022	1.017	0.09115032	4
A5	1.012	0.727	0.527	0.851	1	0.611	0.950	0.869	0.865	0.07752716	8
A6	1.657	1.191	0.862	1.392	1.637	1	1.555	1.423	1.416	0.12691945	2
A7	1.066	0.766	0.555	0.896	1.053	0.643	1	0.915	0.910	0.08162745	7
A8	1.164	0.837	0.606	0.979	1.151	0.703	1.093	1	0.995	0.08919895	6
A9	1.170	0.841	0.609	0.984	1.156	0.706	1.098	1.005	1	0.08965879	5

**Table 14** Pairwise comparison matrix for the alternatives regarding enjoyment in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.427	0.381	0.341	0.573	0.898	1.906	1.301	1.395	0.07020630	6
A2	2.340	1	0.891	0.797	1.340	2.102	4.459	3.045	3.264	0.16427961	3
A3	2.625	1.122	1	0.894	1.503	2.358	5.002	3.416	3.662	0.18429596	2
A4	2.936	1.255	1.119	1	1.681	2.637	5.595	3.820	4.096	0.20614247	1
A5	1.746	0.746	0.665	0.595	1	1.569	3.328	2.272	2.436	0.12260709	4
A6	1.113	0.476	0.424	0.379	0.637	1	2.121	1.449	1.553	0.07815893	5
A7	0.525	0.224	0.200	0.179	0.300	0.471	1	0.683	0.732	0.03684313	9
A8	0.769	0.328	0.293	0.262	0.440	0.690	1.465	1	1.072	0.05395722	7
A9	0.717	0.306	0.273	0.244	0.411	0.644	1.366	0.933	1	0.05033126	8

**Table 15** Pairwise comparison matrix for the alternatives regarding faculty in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	1.754	3.960	2.909	4.457	4.064	3.308	3.830	1.212	0.24214347	1
A2	0.570	1	2.257	1.658	2.541	2.317	1.885	2.183	0.691	0.13802755	3
A3	0.253	0.443	1	0.735	1.125	1.026	0.835	0.967	0.306	0.06114298	7
A4	0.344	0.603	1.361	1	1.532	1.397	1.137	1.317	0.417	0.08324288	4
A5	0.224	0.394	0.889	0.653	1	0.912	0.742	0.859	0.272	0.05432878	9
A6	0.246	0.432	0.974	0.716	1.097	1	0.814	0.942	0.298	0.05957748	8
A7	0.302	0.530	1.197	0.879	1.347	1.229	1	1.158	0.366	0.07320527	5
A8	0.261	0.458	1.034	0.759	1.164	1.061	0.864	1	0.316	0.06321575	6
A9	0.825	1.448	3.269	2.401	3.679	3.355	2.730	3.162	1	0.19985944	2

**Table 16** Pairwise comparison matrix for the alternatives regarding interaction in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.247	0.454	0.515	0.660	0.927	1.536	1.331	1.771	0.07088712	6
A2	4.055	1	1.842	2.088	2.678	3.758	6.229	5.398	7.183	0.28743492	1
A3	2.201	0.543	1	1.134	1.454	2.040	3.381	2.930	3.899	0.15602703	2
A4	1.942	0.479	0.882	1	1.282	1.800	2.983	2.585	3.440	0.13764311	3
A5	1.514	0.373	0.688	0.780	1	1.403	2.326	2.016	2.682	0.10733931	4
A6	1.079	0.266	0.490	0.556	0.713	1	1.658	1.436	1.911	0.07648407	5
A7	0.651	0.161	0.296	0.335	0.430	0.603	1	0.867	1.153	0.04614353	8
A8	0.751	0.185	0.341	0.387	0.496	0.696	1.154	1	1.331	0.05325129	7
A9	0.564	0.139	0.256	0.291	0.373	0.523	0.867	0.751	1	0.04001535	9

**Table 17** Pairwise comparison matrix for the alternatives regarding structure in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.302	0.528	0.505	0.619	1.092	0.376	1.079	1.387	0.06318701	6
A2	3.307	1	1.748	1.670	2.048	3.612	1.242	3.567	4.587	0.20895735	1
A3	1.892	0.572	1	0.956	1.172	2.067	0.711	2.041	2.625	0.11956994	4
A4	1.980	0.599	1.046	1	1.226	2.163	0.744	2.136	2.747	0.12511872	3
A5	1.615	0.488	0.853	0.816	1	1.764	0.607	1.742	2.240	0.10203971	5
A6	0.916	0.277	0.484	0.462	0.567	1	0.344	0.987	1.270	0.05784910	8
A7	2.662	0.805	1.407	1.344	1.649	2.908	1	2.871	3.693	0.16821623	2
A8	0.927	0.280	0.490	0.468	0.574	1.013	0.348	1	1.286	0.05858617	7
A9	0.721	0.218	0.381	0.364	0.446	0.787	0.271	0.778	1	0.04555374	9

**Table 18** Pairwise comparison matrix for the alternatives regarding usefulness in Art, Design, and Humanity course selection

	A1	A2	A3	A4	A5	A6	A7	A8	A9	Priorities	Rankings
A1	1	0.842	0.246	0.654	0.458	0.686	0.769	0.377	0.536	0.04484990	9
A2	1.187	1	0.292	0.776	0.543	0.814	0.913	0.448	0.636	0.05324379	8
A3	4.059	3.419	1	2.655	1.857	2.784	3.122	1.530	2.175	0.18203878	1
A4	1.529	1.288	0.377	1	0.700	1.049	1.176	0.577	0.819	0.06857601	5
A5	2.186	1.841	0.539	1.429	1	1.499	1.681	0.824	1.171	0.09802823	3
A6	1.458	1.228	0.359	0.953	0.667	1	1.121	0.550	0.781	0.06537621	6
A7	1.300	1.095	0.320	0.850	0.595	0.892	1	0.490	0.697	0.05830051	7
A8	2.652	2.234	0.653	1.734	1.213	1.819	2.040	1	1.421	0.11894135	2
A9	1.866	1.572	0.460	1.220	0.854	1.280	1.435	0.704	1	0.08368541	4

**Table 19** Pairwise comparison matrix for the alternatives regarding structure in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	0.879	0.252	0.839	0.824	1.107	0.564	0.363	0.830	0.06603153	8
B2	1.138	1	0.287	0.955	0.938	1.260	0.641	0.414	0.944	0.07515078	7
B3	3.967	3.486	1	3.330	3.271	4.392	2.236	1.442	3.291	0.26197673	1
B4	1.191	1.047	0.300	1	0.982	1.319	0.671	0.433	0.988	0.07866349	6
B5	1.213	1.066	0.306	1.018	1	1.343	0.684	0.441	1.006	0.08009194	4
B6	0.903	0.794	0.228	0.758	0.745	1	0.509	0.328	0.749	0.05964629	9
B7	1.775	1.559	0.447	1.490	1.463	1.965	1	0.645	1.472	0.11717687	3
B8	2.751	2.417	0.693	2.309	2.268	3.046	1.550	1	2.283	0.18167006	2
B9	1.205	1.059	0.304	1.012	0.994	1.334	0.679	0.438	1	0.07959232	5

**Table 20** Pairwise comparison matrix for the alternatives regarding faculty in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	0.657	0.253	0.550	0.688	1.129	0.318	0.306	0.785	0.05460495	8
B2	1.523	1	0.386	0.838	1.048	1.719	0.484	0.467	1.195	0.08315478	5
B3	3.947	2.592	1	2.173	2.717	4.455	1.254	1.209	3.097	0.21554366	1
B4	1.817	1.193	0.460	1	1.251	2.050	0.577	0.557	1.425	0.09920132	4
B5	1.453	0.954	0.368	0.800	1	1.639	0.461	0.445	1.140	0.07931976	6
B6	0.886	0.582	0.224	0.488	0.610	1	0.281	0.271	0.695	0.04838699	9
B7	3.149	2.068	0.798	1.733	2.168	3.554	1	0.965	2.471	0.17195107	3
B8	3.264	2.143	0.827	1.797	2.247	3.684	1.037	1	2.561	0.17823778	2
B9	1.275	0.837	0.323	0.702	0.877	1.438	0.405	0.390	1	0.06959969	7

**Table 21** Pairwise comparison matrix for the alternatives regarding assessment in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	1.007	0.281	1.051	1.153	1.457	0.557	0.581	0.989	0.07950892	5
B2	0.993	1	0.280	1.043	1.145	1.447	0.553	0.577	0.982	0.07895288	6
B3	3.553	3.578	1	3.732	4.096	5.178	1.980	2.064	3.514	0.28247013	1
B4	0.952	0.959	0.268	1	1.097	1.387	0.531	0.553	0.942	0.07568302	7
B5	0.867	0.873	0.244	0.911	1	1.264	0.483	0.504	0.858	0.06896232	8
B6	0.686	0.691	0.193	0.721	0.791	1	0.382	0.399	0.679	0.05455261	9
B7	1.794	1.807	0.505	1.885	2.069	2.615	1	1.042	1.775	0.14265084	2
B8	1.721	1.733	0.484	1.808	1.984	2.508	0.959	1	1.703	0.13684519	3
B9	1.011	1.018	0.285	1.062	1.165	1.473	0.563	0.587	1	0.08037410	4

**Table 22** Pairwise comparison matrix for the alternatives regarding content in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	1.077	0.321	0.977	1.229	1.455	0.653	0.572	0.957	0.08412255	6
B2	0.928	1	0.298	0.907	1.141	1.350	0.606	0.531	0.888	0.07808652	7
B3	3.111	3.351	1	3.039	3.823	4.524	2.031	1.780	2.976	0.26167295	1
B4	1.023	1.103	0.329	1	1.258	1.489	0.668	0.586	0.979	0.08609547	5
B5	0.814	0.877	0.262	0.795	1	1.183	0.531	0.466	0.778	0.06844490	8
B6	0.688	0.741	0.221	0.672	0.845	1	0.449	0.394	0.658	0.05783515	9
B7	1.532	1.650	0.492	1.496	1.882	2.228	1	0.877	1.465	0.12883743	3
B8	1.747	1.882	0.562	1.707	2.147	2.541	1.141	1	1.671	0.14696976	2
B9	1.045	1.126	0.336	1.021	1.285	1.520	0.683	0.598	1	0.08793528	4

**Table 23** Pairwise comparison matrix for the alternatives regarding interaction in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	0.687	0.201	0.758	0.713	1.085	0.426	0.391	0.804	0.05801926	8
B2	1.456	1	0.292	1.105	1.038	1.581	0.620	0.569	1.171	0.08450420	4
B3	4.985	3.423	1	3.781	3.552	5.411	2.123	1.948	4.006	0.28922762	1
B4	1.318	0.905	0.264	1	0.939	1.431	0.561	0.515	1.060	0.07649307	6
B5	1.403	0.964	0.282	1.065	1	1.523	0.598	0.549	1.128	0.08142826	5
B6	0.921	0.633	0.185	0.699	0.656	1	0.392	0.360	0.740	0.05345036	9
B7	2.348	1.612	0.471	1.781	1.673	2.549	1	0.918	1.887	0.13623914	3
B8	2.559	1.757	0.513	1.941	1.823	2.777	1.090	1	2.056	0.14844739	2
B9	1.244	0.854	0.250	0.944	0.887	1.351	0.530	0.486	1	0.07219070	7

**Table 24** Pairwise comparison matrix for the alternatives regarding usefulness in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	0.815	0.252	0.941	0.807	1.145	0.503	0.361	0.831	0.06522387	8
B2	1.227	1	0.309	1.155	0.990	1.404	0.617	0.443	1.019	0.08001422	5
B3	3.965	3.232	1	3.732	3.198	4.538	1.993	1.430	3.294	0.25859129	1
B4	1.062	0.866	0.268	1	0.857	1.216	0.534	0.383	0.883	0.06929152	7
B5	1.240	1.011	0.313	1.167	1	1.419	0.623	0.447	1.030	0.08085850	4
B6	0.874	0.712	0.220	0.822	0.705	1	0.439	0.315	0.726	0.05698581	9
B7	1.989	1.622	0.502	1.872	1.605	2.277	1	0.718	1.653	0.12974468	3
B8	2.772	2.259	0.699	2.609	2.236	3.172	1.393	1	2.303	0.18078532	2
B9	1.204	0.981	0.304	1.133	0.971	1.378	0.605	0.434	1	0.07850481	6

**Table 25** Pairwise comparison matrix for the alternatives regarding effort in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	1.348	0.334	1.386	0.929	1.349	0.707	0.550	1.297	0.08867243	5
B2	0.742	1	0.248	1.028	0.689	1.001	0.524	0.408	0.962	0.06576027	7
B3	2.992	4.035	1	4.148	2.780	4.038	2.114	1.646	3.882	0.26534916	1
B4	0.721	0.973	0.241	1	0.670	0.973	0.510	0.397	0.936	0.06396742	9
B5	1.076	1.451	0.360	1.492	1	1.452	0.760	0.592	1.396	0.09543809	4
B6	0.741	0.999	0.248	1.027	0.689	1	0.524	0.408	0.961	0.06570977	8
B7	1.415	1.908	0.473	1.962	1.315	1.910	1	0.778	1.836	0.12549462	3
B8	1.818	2.452	0.608	2.521	1.690	2.454	1.285	1	2.359	0.16124730	2
B9	0.771	1.040	0.258	1.069	0.716	1.040	0.545	0.424	1	0.06836092	6

**Table 26** Pairwise comparison matrix for the alternatives regarding technicality in Computer Science, Engineering, and Programming course selection

	B1	B2	B3	B4	B5	B6	B7	B8	B9	Priorities	Rankings
B1	1	1.098	0.257	1.138	0.973	1.296	0.665	0.548	0.988	0.07803783	6
B2	0.911	1	0.234	1.037	0.886	1.180	0.605	0.499	0.899	0.07105562	7
B3	3.885	4.267	1	4.423	3.781	5.036	2.584	2.129	3.838	0.30319182	1
B4	0.878	0.965	0.226	1	0.855	1.139	0.584	0.481	0.868	0.06854996	8
B5	1.027	1.128	0.264	1.170	1	1.332	0.683	0.563	1.015	0.08018378	4
B6	0.771	0.847	0.199	0.878	0.751	1	0.513	0.423	0.762	0.06020207	9
B7	1.504	1.652	0.387	1.712	1.464	1.949	1	0.824	1.485	0.11735581	3
B8	1.825	2.004	0.470	2.078	1.776	2.366	1.214	1	1.803	0.14242091	2
B9	1.012	1.112	0.261	1.152	0.985	1.312	0.673	0.555	1	0.07900220	5

**Author contributions** X.C.: Conceptualization, Formal analysis, Data curation, Writing—original draft, Writing—review and editing, Visualization. H.X.: Conceptualization, Methodology, Writing—review and editing, Supervision, Funding acquisition. X.T.: Methodology, Data curation, Conceptualization. F.L.W.: Resources, Supervision, Funding acquisition, Project administration. J.C.: Visualization, Validation, Supervision.

**Funding** This work has been supported by the National Natural Science Foundation of China (No. 62307010), as well as the Direct Grant (DR23B2) and the Faculty Research Grants (DB23A3 and DB23B2) of Lingnan University, Hong Kong.

**Availability of data and materials** The data are available upon reasonable request from the corresponding author.

## Declarations

**Conflict of interest** All authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Nie Y, Luo H, Sun D (2021) Design and validation of a diagnostic MOOC evaluation method combining AHP and text mining algorithms. *Interact Learn Environ* 29:315–328. <https://doi.org/10.1080/10494820.2020.1802298>
- Chen X, Zou D, Cheng G, Xie H (2023) Deep neural networks for the automatic understanding of the semantic content of online course reviews. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-023-11980-6>
- Wei X, Saab N, Admiraal W (2023) Do learners share the same perceived learning outcomes in MOOCs? Identifying the role of motivation, perceived learning support, learning engagement, and self-regulated learning strategies. *Internet High Educ* 56:100880. <https://doi.org/10.1016/j.iheduc.2022.100880>
- Gong J, Wan Y, Liu Y, Li X, Zhao Y, Wang C, Lin Y, Fang X, Feng W, Zhang J (2023) Reinforced moocs concept recommendation in heterogeneous information networks. *ACM Trans Web* 17:1–27. <https://doi.org/10.1145/3580510>
- Gong J, Wang S, Wang J, Feng W, Peng H, Tang J, Yu PS (2020) Attentional graph convolutional networks for knowledge concept recommendation in moocs in a heterogeneous view. In: *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*, Xi'an, China, pp 79–88. <https://doi.org/10.1145/3397271.3401057>
- Luo R, Ye Z (2021) What makes a good-quality language MOOC? An empirical study of criteria to evaluate the quality of online language courses from learners' perspectives. *ReCALL* 33:177–192. <https://doi.org/10.1017/S0958344021000082>
- Zhu M, Berri S, Koda R, Wu Y (2023) Exploring students' self-directed learning strategies and satisfaction in online learning. *Educ Inf Technol*. <https://doi.org/10.1007/s10639-023-11914-2>



8. Albelbisi NA, Al-Adwan AS, Habibi A (2021) Self-regulated learning and satisfaction: a key determinants of MOOC success. *Educ Inf Technol* 26:3459–3481. <https://doi.org/10.1007/s10639-020-10404-z>
9. Dai HM, Teo T, Rappa NA, Huang F (2020) Explaining Chinese university students' continuance learning intention in the MOOC setting: a modified expectation confirmation model perspective. *Comput Educ* 150:103850. <https://doi.org/10.1016/j.compedu.2020.103850>
10. Hone KS, El Said GR (2016) Exploring the factors affecting MOOC retention: a survey study. *Comput Educ* 98:157–168. <https://doi.org/10.1016/j.compedu.2016.03.016>
11. Nanda G, Douglas KA, Waller DR, Merzdorf HE, Goldwasser D (2021) Analyzing large collections of open-ended feedback from MOOC learners using LDA topic modeling and qualitative analysis. *IEEE Trans Learn Technol* 14:146–160. <https://doi.org/10.1109/TLT.2021.3064798>
12. Margaryan A, Bianco M, Littlejohn A (2015) Instructional quality of massive open online courses (MOOCs). *Comput Educ* 80:77–83. <https://doi.org/10.1016/j.compedu.2014.08.005>
13. Rong L, Wang L, Liu P, Zhu B (2021) Evaluation of MOOCs based on multigranular unbalanced hesitant fuzzy linguistic MABAC method. *Int J Intell Syst* 36:5670–5713. <https://doi.org/10.1002/int.22526>
14. Li L, Johnson J, Aarhus W, Shah D (2022) Key factors in MOOC pedagogy based on NLP sentiment analysis of learner reviews: what makes a hit. *Comput Educ* 176:104354. <https://doi.org/10.1016/j.compedu.2021.104354>
15. Nilashi M, Abumalloh RA, Zibarzani M, Samad S, Zogaan WA, Ismail MY, Mohd S, Akib NAM (2022) What factors influence students satisfaction in massive open online courses? Findings from user-generated content using educational data mining. *Educ Inf Technol* 37:9401–9435. <https://doi.org/10.1007/s10639-022-10997-7>
16. Peng X, Xu Q (2020) Investigating learners' behaviors and discourse content in MOOC course reviews. *Comput Educ* 143:103673. <https://doi.org/10.1016/j.compedu.2019.103673>
17. Liu S, Peng X, Cheng HNH, Liu Z, Sun J, Yang C (2019) Unfolding sentimental and behavioral tendencies of learners' concerned topics from course reviews in a MOOC. *J Educ Comput Res* 57:670–696. <https://doi.org/10.1177/0735633118757181>
18. Gomez MJ, Calderón M, Sánchez V, Clemente FJG, Ruipérez-Valiente JA (2022) Large scale analysis of open MOOC reviews to support learners' course selection. *Expert Syst Appl* 210:118400. <https://doi.org/10.1016/j.eswa.2022.118400>
19. Saaty TL (1980) *The analytic hierarchy process*. McGrawhill International, New York
20. Fang J, Partovi FY (2021) Criteria determination of analytic hierarchy process using a topic model. *Expert Syst Appl* 169:114306. <https://doi.org/10.1016/j.eswa.2020.114306>
21. O'Mahony MP, Smyth B (2009) Learning to recommend helpful hotel reviews. In: *Proceedings of the third ACM conference on recommender systems*, New York, United States, pp 305–308. <https://doi.org/10.1145/1639714.1639774>
22. Salton G (1991) Developments in automatic text retrieval. *Science* 253:974–980. <https://doi.org/10.1126/science.253.5023.974>
23. Manning CD (2009) *An introduction to information retrieval*. Cambridge University Press, Cambridge
24. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, Albertson B, Rand DG (2014) Structural topic models for open-ended survey responses. *Am J Pol Sci* 58:1064–1082. <https://doi.org/10.1111/ajps.12103>
25. Blei DM (2012) Probabilistic topic models. *Commun ACM* 55:77–84. <https://doi.org/10.1145/2133806.2133826>
26. Hu N, Zhang T, Gao B, Bose I (2019) What do hotel customers complain about? Text analysis using structural topic model. *Tour Manag* 72:417–426. <https://doi.org/10.1016/j.tourman.2019.01.002>
27. Mimno D, Wallach HM, Talley E, Leenders M, McCallum A (2011) Optimizing semantic coherence in topic models. In: *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics, Edinburgh, Scotland, UK, pp 262–272. <https://aclanthology.org/D11-1024>
28. Lau JH, Newman D, Baldwin T (2014) Machine reading tea leaves: automatically evaluating topic coherence and topic model quality. In: *Proceedings of the 14th conference of the european chapter of the association for computational linguistics*, Gothenburg, Sweden, pp 530–539. <https://doi.org/10.3115/v1/E14-1056>
29. Roberts ME, Stewart BM, Tingley D (2019) Stm: an R package for structural topic models. *J Stat Softw* 91:1–40. <https://doi.org/10.18637/jss.v091.i02>
30. Bischof J, Airoidi EM (2012) Summarizing topical content with word frequency and exclusivity. In: *Proceedings of the 29th international conference on machine learning*, Edinburgh, Scotland, UK, pp 201–208. <https://doi.org/10.48550/arXiv.1206.4631>
31. Chang J, Gerrish S, Wang C, Boyd-Graber J, Blei D (2009) Reading tea leaves: how humans interpret topic models. In: *Proceedings of the 22nd international conference on neural information processing systems*, Vancouver, British Columbia, Canada, pp. 288–296. <https://doi.org/10.5555/2984093.2984126>. Accessed 26 Aug 2023
32. Hew KF, Hu X, Qiao C, Tang Y (2020) What predicts student satisfaction with MOOCs: a gradient boosting trees supervised machine learning and sentiment analysis approach. *Comput Educ* 145:103724. <https://doi.org/10.1016/j.compedu.2019.103724>
33. Du Y, Gao H (2022) Determinants affecting teachers' adoption of AI-based applications in EFL context: An analysis of analytic hierarchy process. *Educ Inf Technol* 27:9357–9384. <https://doi.org/10.1007/s10639-022-11001-y>
34. Chen X, Zou D, Cheng G, Xie H (2020) Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: a retrospective of all volumes of computer & education. *Comput Educ* 151:103855. <https://doi.org/10.1016/j.compedu.2020.103855>
35. Chen X, Zou D, Xie H, Wang FL (2023) Metaverse in education: contributors, cooperations, and research themes. *IEEE Trans Learn Technol* 16:1111–1129. <https://doi.org/10.1109/TLT.2023.3277952>
36. Al-Rahmi WM, Yahaya N, Alamri MM, Alyoussef IY, Al-Rahmi AM, Kamin YB (2021) Integrating innovation diffusion theory with technology acceptance model: supporting students' attitude towards using a massive open online courses (MOOCs) systems. *Interact Learn Environ* 29:1380–1392. <https://doi.org/10.1080/10494820.2019.1629599>
37. Aparicio M, Oliveira T, Bacao F, Painho M (2019) Gamification: a key determinant of massive open online course (MOOC) success. *Inf Manag* 56:39–54. <https://doi.org/10.1016/j.im.2018.06.003>
38. Kim H-W, Chan HC, Gupta S (2007) Value-based adoption of mobile internet: an empirical investigation. *Decis Support Syst* 43:111–126. <https://doi.org/10.1016/j.dss.2005.05.009>
39. Sun Y, Ni L, Zhao Y et al (2019) Understanding students' engagement in MOOCs: an integration of self-determination theory and theory of relationship quality. *Br J Educ Technol* 50:3156–3174. <https://doi.org/10.1111/bjet.12724>
40. Ramesh A, Kumar SH, Foulds J, Getoor L (2015) Weakly supervised models of aspect-sentiment for online course discussion forums. In: *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint*

- conference on natural language processing, Beijing, China, pp 74–83. <https://doi.org/10.3115/v1/P15-1008>
41. Hagerer G, Lahesoo L, Anschütz M, Krusche S, Groh G (2021) An analysis of programming course evaluations before and after the introduction of an autograder. In: 2021 19th international conference on information technology based higher education and training (ITHET), Sydney, Australia, pp 1–9. <https://doi.org/10.48550/arXiv.2110.15134>
  42. Reif F (1987) Interpretation of scientific or mathematical concepts: cognitive issues and instructional implications. *Cogn Sci* 11:395–416. [https://doi.org/10.1016/S0364-0213\(87\)80014-9](https://doi.org/10.1016/S0364-0213(87)80014-9)
  43. Hew KF (2016) Promoting engagement in online courses: what strategies can we learn from three highly rated MOOCs. *Br J Educ Technol* 47:320–341. <https://doi.org/10.1111/bjet.12235>
  44. Watson SL, Watson WR, Janakiraman S, Richardson J (2017) A team of instructors' use of social presence, teaching presence, and attitudinal dissonance strategies: an animal behaviour and welfare MOOC. *Int Rev Res Open Distrib Learn* 18:68–91. <https://doi.org/10.19173/irrodl.v18i2.2663>
  45. Swan K, Garrison DR, Richardson JC (2009) A constructivist approach to online learning: the community of inquiry framework. In: Payne CR (ed) *Information technology and constructivism in higher education: progressive learning frameworks*. IGI Global, New York, pp 43–57. <https://doi.org/10.4018/978-1-60566-654-9.ch004>
  46. Lim JRN, Rosenthal S, Sim YJM et al (2021) Making online learning more satisfying: the effects of online-learning self-efficacy, social presence and content structure. *Technol Pedagog Educ* 30:543–556. <https://doi.org/10.1080/1475939X.2021.1934102>
  47. Oyarzun B, Barreto D, Conklin S (2018) Instructor social presence effects on learner social presence, achievement, and satisfaction. *TechTrends* 62:625–634. <https://doi.org/10.1007/s11528-018-0299-0>
  48. Li N, Kidziński Ł, Jermann P, Dillenbourg P (2015) MOOC video interaction patterns: What do they tell us? In: *European conference on technology enhanced learning*, Toledo, Spain, pp 197–210. [https://doi.org/10.1007/978-3-319-24258-3\\_15](https://doi.org/10.1007/978-3-319-24258-3_15)
  49. Howell GF, Buck JM (2012) The adult student and course satisfaction: What matters most? *Innov High Educ* 37:215–226. <https://doi.org/10.1007/s10755-011-9201-0>
  50. Doo MY, Tang Y, Bonk CJ, Zhu M (2020) MOOC instructor motivation and career development. *Distance Educ* 41:26–47. <https://doi.org/10.1080/01587919.2020.1724770>
  51. Wu B, Chen W (2023) Factors affecting MOOC teacher effectiveness from the perspective of professional capital. *Behav Inf Technol* 42:498–513. <https://doi.org/10.1080/0144929X.2021.2024596>
  52. Tang H, Dai M, Yang S, Du X, Hung JL, Li H (2022) Using multimodal analytics to systemically investigate online collaborative problem-solving. *Distance Educ* 43:290–317. <https://doi.org/10.1080/01587919.2022.2064824>
  53. Chang C-C, Wang Y-H (2021) Using phenomenological methodology with thematic analysis to examine and reflect on commonalities of instructors' experiences in MOOCs. *Educ Sci* 11:203. <https://doi.org/10.3390/educsci11050203>
  54. Sungkur RK, Maharaj MS (2021) Design and implementation of a SMART learning environment for the upskilling of cybersecurity professionals in Mauritius. *Educ Inf Technol* 26:3175–3201. <https://doi.org/10.1007/s10639-020-10408-9>
  55. Bali M (2014) MOOC pedagogy: glean good practice from existing MOOCs. *J Online Learn Teach* 10:44. [https://jolt.merlot.org/vol10no1/bali\\_0314.pdf](https://jolt.merlot.org/vol10no1/bali_0314.pdf). Accessed 26 Aug 2023
  56. Zimmerman BJ (2008) Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *Am Educ Res J* 45:166–183. <https://doi.org/10.3102/0002831207312909>
  57. Deshpande A, Chukhlomin V (2017) What makes a good MOOC: a field study of factors impacting student motivation to learn. *Am J Distance Educ* 31:275–293. <https://doi.org/10.1080/08923647.2017.1377513>
  58. Bey A, Jermann P, Dillenbourg P (2018) A comparison between two automatic assessment approaches for programming: an empirical study on MOOCs. *J Educ Technol Soc* 21:259–272
  59. Davis FD, Bagozzi RP, Warshaw PR (1989) User acceptance of computer technology: a comparison of two theoretical models. *Manag Sci* 35:982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
  60. Davis FD, Bagozzi RP, Warshaw PR (1992) Extrinsic and intrinsic motivation to use computers in the workplace. *J Appl Soc Psychol* 22:1111–1132. <https://doi.org/10.1111/j.1559-1816.1992.tb00945.x>
  61. Oliver RL (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *J Mark Res* 17:460–469. <https://doi.org/10.1177/002224378001700405>
  62. Howarth JP, D'Alessandro S, Johnson L, White L (2016) Learner motivation for MOOC registration and the role of MOOCs as a university 'taster.' *Int J Lifelong Educ* 35:74–85. <https://doi.org/10.1080/02601370.2015.1122667>
  63. Christensen G, Steinmetz A, Alcorn B, Bennett A, Woods D, Emanuel E (2013) The MOOC phenomenon: Who takes massive open online courses and why? Available SSRN 2350964. <https://doi.org/10.2139/ssrn.2350964>
  64. Bayeck R (2016) Exploratory study of MOOC learners' demographics and motivation: the case of students involved in groups. *Open Prax* 8:223–233. <https://doi.org/10.5944/openpraxis.8.3.282>
  65. Principe VA, de Souza Vale RG, de Castro JBP, Carvano LM, Henriques RAP, de Almeida e Sousa Lobo VJ, de Alkmim Moreira Nunes R (2022) A computational literature review of football performance analysis through probabilistic topic modeling. *Artif Intell Rev* 55:1351–1371. <https://doi.org/10.1007/s10462-021-09998-8>
  66. Muralidhara S, Paul MJ (2018) # Healthy selfies: exploration of health topics on Instagram. *JMIR Public Heal Surveill* 4:e10150. <https://doi.org/10.2196/10150>
  67. Gregoriades A, Pampaka M, Herodotou H, Christodoulou E (2021) Supporting digital content marketing and messaging through topic modelling and decision trees. *Expert Syst Appl* 184:115546. <https://doi.org/10.1016/j.eswa.2021.115546>
  68. Kraus S, Kumar S, Lim WM, Kaur J, Sharma A, Schiavone F (2023) From moon landing to metaverse: tracing the evolution of Technological Forecasting and Social Change. *Technol Forecast Soc Change* 189:122381. <https://doi.org/10.1016/j.techfore.2023.122381>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.