Learning Behavior Analysis of MOOC Learners Based on Multivariate Meta-Analysis Model

Yang Le
College of Computer and Information Engineering, Jiangxi Agriculture University, Nanchang, China

Keywords: Multivariate Meta-Analysis Model, MOOC, Behaviour Model, Learning Effectiveness, Influence.

Abstract: MOOC is promoting another innovation in human education, which realizes a broad increase in educational participation. MOOC supports and facilitates learners' autonomous personalized learning process and learning outcomes, and MOOC learners have diversified learning behaviour patterns. Many studies have pointed out that online learning behaviour patterns are highly correlated with learning outcomes. This study proposes a learning behaviour analysis architecture for MOOC learners based on multivariate meta-analysis, taking the "Web System Development and Design" course as an example, online learning behaviour data of the MOOC learners are processed and analysed to obtain the desired result effect value, and merging all the effect values by multivariate analysis, finally obtaining a valid conclusion. According to the results of data analysis, relevant improvement strategies and suggestions are put forward for different stakeholders. By discovering the problems hidden in the online learning process of MOOC learners, prompting and warnings in a timely manner, and adjusting the teaching plan in a targeted manner.

1 INTRODUCTION

Since the outbreak of novel coronavirus pneumonia (COVID-19) at the end of 2019, the development of online learning has been greatly accelerated. As a new form of education, MOOC has attracted much attention for its large-scale, open, free learning platform services and other characteristics. It transcends the limitation of time and space and realizes the wide participation of education. In the context of "Internet +", big data is rapidly flooding all walks of life, and the field of education is no exception. With the development of learning analysis and artificial intelligence technology, it has become an irresistible trend to use online learning behaviour data of student to evaluate students' learning performance and provide better services for students (Francisco, 2018; Tsai 2018; Gold 2020). The online learning platform will record the basic information of MOOC learners, the update of teaching resources and the learning behaviour data of learners, that is, the online learning platform stores the learning behaviour data of MOOC learners from the beginning to the end of learning. In the face of massive data, how to collect, process and analyse the data so as to truly understand MOOC learners' course learning situation, value orientation and psychological state of MOOC learners, making MOOC teaching more attractive and spreadable, is one of the key tasks to be completed urgently to discover and use the value of online learning behaviour data, and to provide strong data support, method compliance and practical guidance for promoting educational informatization to help families, schools and society cooperate in educating people

Jovanovic et al. used the K-means algorithm to construct evaluation criteria to distinguish learners into three categories: good, general and poor. On this basis, they continued to use the cognitive characteristics of learners collected by the scale and made further analysis and exploration (Jovanovic 2012). Using the back propagation (BP) neural network method, Wang et al. developed an English learning system that matches the learner 's learning situation. This system could recommend learning resources suitable for each learner based on different learner's personality traits, gender, and learner learning anxiety (Wang 2011). Rajendran et al. comprehensively analyzed the relevant theories of pedagogy and psychology, and they found that if learners encounter difficulties in the process of learning, the learners will be frustrated. Then, they

Le, Y.
Learning Behavior Analysis of MOOC Learners Based on Multivariate Meta-Analysis Model.
DOI: 10.5220/0011901400003613
In *Proceedings of the 2nd International Conference on New Media Development and Modernized Education (NMDME 2022)*, pages 130-136
ISBN: 978-989-758-630-9

analyzed the situation of learners using intelligent auxiliary systems, collected the reasons for their learning difficulties and established a linear model to achieve the goal of judging the emotional state of learners when using intelligent auxiliary systems, and timely reminded teachers to resolve the difficulties encountered by learners and changed the current situation that learners are not conducive to learning (Rajendran 2013). Aher et al. used the K-means algorithm to cluster the basic situation of learners, then, they used the Apriori algorithm to analyze the correlation of learners in each category, and obtained the course categories that learners like to learn, so as to push their favorite learning content to learners (Aher 2013). Chen implemented a multi-label classification algorithm to classify tweets that reflect student issues, selecting about 35,000 tweets from Purdue University to train a problem detector, demonstrating how informal social media data provides insights into student experiences and strategies (Chen 2014). Patil gathered information about all engineering students' online interactions on Twitter, analyzed problems such as heavy learning burdens, negative emotions, lack of social engagement, and drowsiness. At the same time, he used Bayesian algorithms to process the data and tried to solve this problem (Patil 2018). Guo proposed a multiple learning behavior analysis framework. Based on the perspective of multiple frameworks, she systematically analyzed the learning behavior of MOOC learners participating in X course, and discussed the strategies for optimizing the design of MOOC courses (Guo 2017). Shen et al. obtained a lot of MOOC learners' learning behavior data from relevant platforms, constructed a model of students' online learning behavior and online learning performance evaluation, and then they conducted sampling stepwise regression to understand the impact of students ' online learning behavior on their academic performance (Shen 2020). Based on the perspective of network learning resources, Zhao et al. empirically studied the learning behavior pattern of online learners and its influence on learning effectiveness, indicating that the behavior pattern of accessing network learning resources is related to learning effectiveness (Zhao 2019). Li focused on the learning behavior of MOOC learners, explored the impact of MOOC learners' learning behavior patterns on learning effectiveness, and provided effective suggestions for improving the learning effectiveness of MOOC learners (Li 2020). Cheng et al. combined the course of "Principles of Systems Engineering" on the military vocational education platform to collect online learning data for analyzing the learning behavior of MOOC learners in military education, and proposed the optimization method of MOOC design in military education based on learning behavior analysis (Cheng 2022).

These studies have systematically analysed the learning behaviour data of MOOC learners, to a certain extent; can effectively promote the learning completion rate of MOOC courses. But based on the perspective of multivariate meta-analysis to study the learning behaviour of MOOC learners, there is less involved. Therefore, this paper studies the analysis of MOOC learners' online learning behaviour data in the multivariate meta-analysis environment, constructs the analysis framework of MOOC learners' online learning behaviour based on multivariate meta-analysis, and combines chapter learning and video learning to empirically analyze MOOC learners' online learning behaviour, so as to provide operational opinions and suggestions for improving MOOC learners ' learning effectiveness.

2 CONSTRUCT MULTIVARIATE META-ANALYSIS FRAMEWORK FOR MOOC LEARNERS' LEARNING BEHAVIOR

Taking MOOC learners who take the course of "Web System Development and Design" as samples, the system log data of MOOC platform and the basic information of learners are collected to analyse the preliminary influence relationship between MOOC learners' learning behaviour and learning effect. Firstly, the analysis framework of MOOC learners' learning behaviour is designed, and MOOC learners' learning behaviour is divided into five categories: (1) Resource learning behaviour: such as the total number of platform login, the total duration of platform login, the total number of resource learning, the total duration of resource learning, resource learning completion rate, resource learning interval, resource learning hops, whether learners learn resources in order, resource learning repetition rate and so on; (2) Homework learning behaviour: such as job scoring rate, job completion rate, number of repeated jobs, number of repeated submission tests and so on; (3) Interactive learning behaviour: such as the number of MOOC learners browsing posts, the total number of posts, the total number of replies, the number of posts, and the number of replies and so on; (4) Learning time preference behaviour: such as the number and time of morning study, the number and time of afternoon study, the number and time of evening study, the number and time of study in the early hours of the morning and so on; (5) Page access behaviour: such as course announcement access times, courseware access times, course scoring standard access times, other page access times, etc.

Before the empirical analysis framework, this study first used confirmatory factor analysis to evaluate the various latent variables in the research model, and used reliability analysis to measure the adequacy of the model. Cronbach's coefficient (α) and composite reliability (CR) were used to evaluate the reliability of the measurement model. It can be seen from Table 1 that these two indicators are greater than the critical value of 0.70 in each latent variable, indicating that the measurement model of each latent variable has good internal consistency, so the MOOC learner learning behaviour analysis framework of multivariate meta-analysis constructed in this study has good reliability.

Table 1: Reliability analysis of Analysis framework.

Classification	Average	Standard deviation	α	CR
Resource learning behaviour	3.732	0.891	0.882	0.892
Homework learning behaviour	3.043	0.878	0.788	0.853
Interactive learning behaviour	2.987	0.868	0.796	0.782
Learning time preference behaviour	2.961	0.845	0.789	0.799
Page access behaviour	3.661	0.816	0.852	0.815

3 AN EMPIRICAL ANALYSIS OF MOOC LEARNERS' LEARNING BEHAVIOUR

3.1 MOOC Learner Learning Behaviour Data Processing

The collected data of MOOC learners' online learning behaviour are derived from the explicit behaviour of "Web System Development and Design" course. Firstly, all the data are desensitized to ensure the personal privacy of MOOC learners. Part of the MOOC learners' online learning behaviour data collected is the personal basic information of MOOC learners, including name, education, major, etc. The other part is the explicit behaviour data of MOOC learners, including the number of course task points completed, audio and video viewing details, discussion details, chapter learning times, assignments, comprehensive grades, etc. This paper will select the appropriate model through the above data to explore MOOC learners' online learning behaviour and the degree of mastery of the course, understand the relationship between course learning grades, and put forward corresponding conclusions and problem-solving measures for MOOC learners, teachers, teaching designers and teaching managers. Firstly, based on the idea of the distribution lag nonlinear model, the nonlinear exposure-response relationship between each explicit behaviour-performance was fitted respectively. The natural cubic spline function was used to control the confounding of audio and video, discussion, the number of chapters learning and grade. After the effect of single explicit behaviour was obtained, the effect value of the result was merged and analysed by multivariate Meta random effect model.

3.2 MOOC Learner Learning Behaviour Data Processing

There are many data in the analysis framework of MOOC learners' learning behaviour, among which chapter learning and video learning in resource learning behaviour are the most representative, which have the greatest impact on MOOC learners' learning effectiveness. After processing the data of MOOC learners' learning behaviour, the multivariate Meta random effect model is used to fit the relationship between the deviation of chapter learning times and the comprehensive performance. The main task of this stage is to estimate the effect of individual explicit behaviour, construct Poisson regression and correct the underestimation of standard error. And then, the result effect value was combined and analysed by using the random effect model of multivariate meta-analysis after obtaining the result effect value of a single class.

3.2.1 Multivariate Meta-Analysis of the Number of Chapters and Grades

The number of chapters in the resource learning behaviour can reflect the degree of completion and enthusiasm of MOOC learners for course learning. A multivariate meta-analysis of the number of chapters and academic grades is conducted to know the implicit influence between the two, understand the inherent problems of MOOC learners' online learning, and better improve the teaching of MOOC courses. The number of chapters is shown in Table 2. The multivariate meta-analysis model is used to process the compiled chapter learning data, and the results are shown in Table 3 and Table 4.

Table 2 Statistics table of chapter learning times.

	Chapter data
Sample size	8065
Minimum number of learning	1073
Maximum number of learning	4077
Average	2503
Deviation	60

Table 3 Results Estimated effect coefficient of Chapter Learning.

	b1	b2	b3	b4
Estimate	-0.3796	-0.4628	-0.7585	-0.0712
Std. Error	0.0466	0.0447	0.0977	0.0268
z	-8.1390	-10.3612	-7.7617	-2.6595
Pr (>|z|)	0.000*	0.000*	0.000*	0.0078
95%ci.lb	-0.4711	-0.5504	-0.9501	-0.1238
95%ci.ub	-0.2882	-0.3753	-0.5670	-0.0187

Table 4 Variance components of Chapter Learning.

	Std. Dev	Corr		
b1	0.1368	b1	b2	b3
b2	0.1273	1.0000		
b3	0.2787	0.9955	0.9959	
b4	0.0590	0.9029	0.9029	0.8581

Table 5 Multivariate meta-analysis of online learning behaviour data heterogeneity.

Multivariate Cochran Q test for heterogeneity
Q = 98.7166 (df = 36)
p-value = 0.000*
I-square statistic = 63.5%

The results of Table 3 showed that the coefficients of the nonlinear result effect curves of the number of chapter's deviation are statistically significant because the P values are less than 0.001 and the P values of b4 are less than 0.01, which can promote MOOC learners to make further efforts in learning. The heterogeneity analysis results of multivariate meta-analysis of online learning behaviour data are shown in Table 5, where Q=98.7166, heterogeneity size I_2=63.5%, and P<0.001, it is explained that the heterogeneity of the collected MOOC learner online learning behaviour data can analyse whether there is a correlation between different learning behaviour patterns and learning outcomes of MOOC learners, and how correlated they are, and promote the construction of high-quality online learning resources and platforms.

3.2.2 Multivariate Meta-Analysis of Time to View Learning Video and Comprehensive Scores

For the explicit behaviour analysis of MOOC learners' online learning behaviour, an important indicator to judge the degree of course learning completion and learning autonomy is the viewing time of course video. A simple multivariate meta-analysis using correlation fitting models to explore the deeper relationship between the time to view learning video and comprehensive scores is an important part of online learning behaviour analysis. Firstly, the collected MOOC learners' learning behaviour data are summarized. According to the characteristics of course video viewing time, the data are sorted out to make simple descriptive statistic, as shown in Table 6. The results showed that MOOC learners have significant differences in the total duration of course video viewing, and the results of multivariate meta-analysis model are shown in Table 7 and 8.

Table 6 Statistics table of learning video viewing time.

	Video data
Sample size	9165
The shortest viewing time（minute）	48.1
The longest viewing time (minute)	1027.8
Average (minute)	595.5
Deviation (minutes)	159.5

Table 7 Results Estimated effect coefficient of Video learning.

	b1	b2	b3	b4
Estimate	-0.5908	-0.6150	-0.7517	-0.8004
Std. Error	0.0478	0.0381	0.0449	0.0409
z	-12.3522	-16.1320	-16.7346	-19.5799
Pr (>|z|)	0.000*	0.000*	0.000*	0.000*
95%ci.lb	-0.6845	-0.6897	-0.8387	-0.8806
95%ci.ub	-0.4970	-0.5403	-0.6636	-0.7203

Table 8 Variance components of Video learning.

	Std. Dev	Corr		
b1	0.0917	b1	b2	b3
b2	0.0729	.9461		
b3	0.0905	0.9704	0.9963	
b4	0.0652	0.6983	0.8924	0.8505

Table 9 Heterogeneity of online learning behaviour data.

Multivariate Cochran Q test for heterogeneity
Q = 103.2740 (df = 45)
p-value = 0.000***
I-square statistic = 56.4%

The P values in Table 7 are all less than 0.001, indicating that the coefficient of the nonlinear result effect curve of learning video viewing time has important analytical significance for improving the stickiness of MOOC learners. In Table 9, Q = 103.2740, heterogeneity size I_2 = 56.4%, and P < 0.001, indicating that the heterogeneity of students' online learning behaviour data collected can promote the improvement of teaching video, attract MOOC learners to learn courses, and is conducive to the completion rate of MOOC courses.

4 IMPROVEMENT STRATEGIES FOR MOOC LEARNING

According to the analysis framework of MOOC learners' learning behaviour and the empirical analysis experiment of MOOC learners' learning behaviour, the problems faced by ordinary MOOC learners in the process of learning MOOC courses are revealed to some extent. Therefore, from the perspective of different stakeholders, this paper puts forward suggestions and strategies to optimize the construction of MOOC courses and to benefit MOOC learners ' learning.

(1) Diversified construction of MOOC learning resources and balanced promotion of the application of various resources.

The study found that the frequency of MOOC learners using various types of learning resources is: video learning resources, picture learning resources, document learning resources, job resources and interactive behaviour. The frequency of MOOC learners browsing learning resources, the total length of browsing and the total number of times determine the learning efficiency of MOOC learners. Therefore, it is suggested to make multi-dimensional video learning resources to guide MOOC learners to learn video in order and regularly. The balanced development of various course learning resources can promote MOOC learners to use and learn various resources from multiple angles and perspectives.

(2) Build a MOOC learning community and enhance the activity of the platform.

The learning method of 'time and space separation' causes MOOC learners to need more social interaction, and MOOC learners from all over the world gather through the MOOC platform with a common goal. Building a MOOC learning community can narrow the interaction distance of social interaction and achieve effective learning. MOOC learning community refers to a virtual learning group composed of teachers and students in the learning process of MOOC courses. Through mutual help, mutual influence, communication and discussion, cooperation and sharing with other, they experience learning and jointly complete the learning tasks of MOOC courses, which will help to effectively improve the teaching effect of the course. At the same time, we should encourage MOOC learners to carry forward the spirit of 'co-creation and sharing', build multi-participation and multi-win network learning resources on the platform, and jointly build rich and diverse curriculum learning resources.

(3) Constructing diversified homework assessment mode to embody knowledge transferable skills.

Homework learning behaviour has a great positive impact on MOOC learners' learning performance, and the completion rate of homework is obviously conducive to the improvement of final grades. Therefore, in the process of MOOC teaching, we should grasp the key factors that affect the learning effect, that is, the proportion of homework completion and the quality of homework completion. There is a significant gap between this influencing factor and other factors, and it is even more important than other influencing factors. Therefore, in order to fully reflect the professional knowledge and skills of students, it should be decided to abandon the simple scheme of using the final exam results as the final score of the course, and it is necessary to establish a scientific and diversified assessment method to reflect the knowledge transferability of MOOC learners. In the process of MOOC teaching, teachers should pay special attention to the design, submission, evaluation and other aspects of the homework part, and urge and ensure that MOOC learners complete their homework on time and efficiently.

(4) Establishing MOOC teaching community to improve course teaching interaction design.

Good curriculum teaching interaction design effect will significantly affect MOOC learner stickiness, and a diverse teaching team can ensure the quality of curriculum teaching interaction design. The purpose of forming the MOOC Teaching Community is to strengthen cooperation, dialogue and sharing among teachers, find effective practices in the teaching process of MOOC courses, and ensure the learning quality of MOOC learners. Teachers need to frequently review the questions raised by MOOC learners and solve problems encountered in the teaching process of MOOC courses. Teachers should lead MOOC learners to discuss and answer questions in a timely manner, and design more reasonable course teaching interactions based on learner satisfaction (Yang 2020). The teaching community must strengthen the cultivation of teachers' informatization ability and strengthen the teaching design skills of MOOC courses. Teachers can carry out a variety of interactive activities (such as timely response, instant communication, etc.), so that learners get a higher sense of participation and enhance learners ' participation in generative resources.

(5) MOOC instructional design and instructional administrators provide personalized learning support services.

MOOC instructional design and teaching management mechanism need to take some measures to optimize the presentation of learning activities, to provide MOOC learners with more personalized learning support and services, thereby enhancing the effectiveness of MOOC learners. For different types of MOOC learners, the learning situation of MOOC learners is predicted according to MOOC big data, and an autonomous personalized learning trajectory is recommended for MOOC learners. MOOC learners independently select other learning resources according to their own preferences in the learning process, and formulate appropriate learning strategies and learning plans (Yang 2016). In the teaching design of MOOC course, it provides learning resources that adapt to the learning methods and habits of MOOC learners, so as to improve the utilization of MOOC course resources by MOOC learners, so as to improve the learning effect of MOOC learners. The design of learning activities is diversified, and the evaluation method of learning effectiveness is more scientific and reasonable, so as to achieve the goal of enhancing the stickiness of MOOC courses.

5 CONCLUSIONS

MOOC learners with different goals, preferences and learning motivations may show different levels of participation and behaviour patterns in learning, and different levels of participation and behaviour patterns will in turn affect learning effectiveness. This paper studies the data of MOOC learners ' learning behaviour, and conducts an empirical analysis of MOOC learners ' chapter learning and video learning based on the multivariate Meta random effect model to explore learners ' different behaviour patterns and the influence of different behaviour patterns on learning effectiveness. It provides useful information for MOOC instructional designers, and then improves the existing instructional design, teaching content arrangement and online learning resource presentation. At the same time, it can also guide MOOC instructional designers to put forward instructional design problems from the perspective of learning theory, and provide a way to answer questions for instructional design, and then develop a personalized learning support system. Combined with the empirical analysis of the results of MOOC learners ' online learning behaviour, it provides effective strategies for improving the learning effectiveness of MOOC learners, and provides reasonable suggestions for the promotion and development of MOOC in the future.

ACKNOWLEDGEMENTS

This work was supported by the 14th Five-years Plan of Jiangxi science of education the key Research topics (No. 22ZD014).

REFERENCES

Aher S B, Lobo LMRJ. Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data[J]. Knowledge-Based Systems, 2013, 51(10):1-14.

Chen X, Vorvoreanu M, Madhavan K. Mining Social Media Data for Understanding Students' Learning Experiences [J]. IEEE Transactions on Learning Technologies, 2014, 7(3):246-259.

Cheng Zhijun, Peng Juan. Research on the Learning Behavior and the Teaching Design of MOOCs for Military Professional Education [J]. Journal of Higher Education Research, 2022, 45(2):65-71.

Francisco J, García P, Ángel F B, et al. Anadaptive hybrid MOOC model: Disrupting the MOOC concept in higher education [J]. Telematics and Informatics, 2018, 35(4):1018-1030.

Gold R, Hemberg E, O'Reilly U M. Analyzing K-12 Blended MOOC Learning Behaviors [C]. Seventh (2020) ACM Conference on Learning, 2020, 8: 12-14.

Guo Xinyue. Analyzing MOOC Learning Behavior under A Multi-Analysis Framework [J]. Journal of Open Learning, 2017, 22(4):21-28.

Jovanovic M, Vukicevic M, Milovanovic M, et al. Using data mining on student behavior and cognitive style data for improving e-learning systems: a case study [J]. International Journal of Computational Intelligence Systems, 2012, 5(3):597-610.

Li Min. A Study on the Influence of Learning Behavior Patterns on Learning Effectiveness in MOOCs [D]. Wuhan: Central China Normal University, 2020.

Patil S, Kulkarni S. Mining Social Media Data for Understanding Students' Learning Experiences using Memetic algorithm [J]. Materials Today Proceedings, 2018, 5(1):693-699.

Rajendran R, Iyer S, Murthy S, et al. A Theory-Driven Approach to Predict Frustration in an ITS [J]. IEEE Transactions on Learning Technologies, 2013, 6(4):378-388.

Shen Xinyi, Liu Meichen, Wu Jianwei, Dong Xianlei.. Towards a model for evaluating students online learning behaviors and learning performance [J]. Distance education in China, 2020, (10),1-8,76.

Tsai Y H, Lin C H, Hong J C, et al. the effects of metacognition on online learning interest and continuance to learn with MOOCs[J]. Computers & Education, 2018, 121:18-29.

Wang Y H, Liao H C.Data mining for adaptive learning in a TESL-based e-learning system [J]. Expert Systems with Applications, 2011,38(6):6480-6485.

Yang le. The Model of Social Interactions 'Influence on the Stickiness of MOOC Learners [J]. Journal of Jiangxi Radio & TV University, 2020,9:19-24.

Yang le. Modeling Personalized and Self-regulated in MOOC [J]. Contemporary Educational Science, 2016(5):40-42.

Zhao Chengling, Li Min, Shu Fengfang, Huang Yan. Research on Models of Participation in Online Learning Resources and Learning Effectiveness [J]. Modern Distance Education, 2019(4):20-27.