Application Research of an Innovative Online Education Model in Big Data Environment

Shi-Yong Zheng
Wuhan University, Hubei, China

Su-Ping Jiang (✉)
Guangxi Talent International College, Qinzhou Guangxi, China
39229374@qq.com

Xiao-Guang Yue
Rajamangala University of Technology, Rattanakosin, Thailand

Ruihui Pu
National Institute of Development Administration, Bangkok, Thailand

Bi-Qing Li
Hezhou University, Hezhou Guangxi, China

Abstract—Online education is a network-based approach to teaching. It is a method of content dissemination and rapid learning through the application of UGC and Internet technology. Compared with traditional school education, online learning can obtain more resources, more autonomy, and no longer limited time and space for learning. Through the questionnaire, this paper finds that learners in the online education model still have some shortcomings in the learning process. For example, the learning process is not durable. Therefore, this paper uses neural network classification algorithm to analyze the related factors that affect the learning behavior of online education students, and propose corresponding control strategies for different influencing factors. By constructing a learning process control strategy model for large educational data, to help learners improve their learning efficiency, help the online education model break through the bottleneck, the online education industry has maintained rapid development. Finally, through the comparative analysis of the improved online education model and the traditional online education model, finding an improved online education model can better improve students' interest in learning. Provided a reference for the development of online education, it also provides a reference for the transformation and upgrading of traditional education to online education.

1 Introduction

The upgrade of Internet era from 1.0 to 2.0 has realized the transformation from information sharing to user social interaction, accompanying by the change of Information content from enterprise creation to user creation (UGC). The characteristics of mobile Internet era are described as SoLoMo Social/Location/mobility. Online education as a new type of education has brought a lot of convenience to users' life and study through social media. It allows users not only to learn knowledge from online education anytime and anywhere, but also to share excellent teaching resources, making study time-saving and efficient. Online education manifests the diffusion of information/knowledge in the social network in the era of "Internet+". All users (including students and teachers), as the creators and disseminators of information, gather on a variety of online education platforms and form information-sharing communities.

In this regard, online education platform can be considered as a kind of virtual community with the purpose of information sharing/interaction, in which the dissemination of knowledge/information can be accomplished by a small number of seed users (opinion leaders). In other words, by empowering only these seed users, knowledge/information can be disseminated rapidly through their social relations in virtual communities. New knowledge will be updated in the context of swarm intelligence. The process of knowledge e-diffusion on the online education platform is presented in the following figure1.
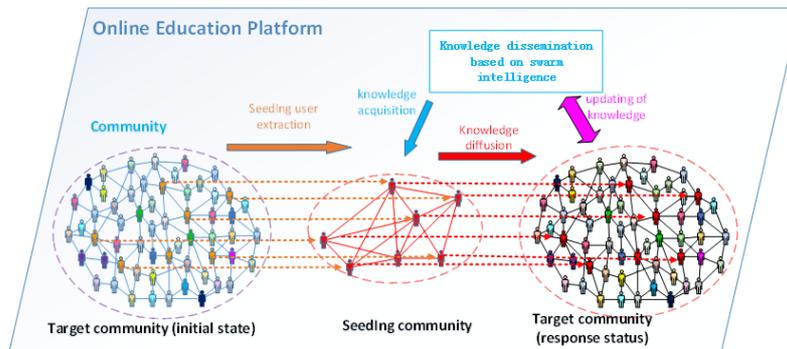

Fig. 1. The Processing of Knowledge Diffusion in Online Education

Online education breaks the boundaries of time, space, age, and virtual and reality, and it is extremely open. On the one hand, the development of online education has made various high-quality educational resources no longer belong to a certain region, a certain school, or a certain city. It has broken through the space constraints, so that students from all over the country and the world can access excellent ones. Teachers, enjoying high-quality teaching achievements and accepting advanced teaching ideas, thus changing the closedness of education and promoting the circulation of knowledge; on the other hand, traditional school education has strict schooling time

and limits student learning activities to a certain extent. Inside, while online education, learners can learn at any time and place according to their own situation. Online education, both in terms of resources and learners, reflects its openness and inclusiveness, which is difficult to achieve in traditional offline education. At present, there are various online education platforms in the online education market, and each online education also provides students with a variety of courses. Students can freely choose an online education platform or course based on their needs. In traditional school education, the progress of learning is determined by the school and the teacher. What tasks need to be completed during each lesson, and what assignments are made after class, are not determined by the students. The students simply obey. Through online education, students need to plan their own learning activities. In the process of learning, students do not have the supervision and supervision of teachers in traditional education. Students must rely on their own will to supervise and control themselves, which greatly tests the students' autonomy ability. The biggest difference between online education and traditional online education is that online education can realize the barrier-free interaction between students and students, teachers and students, platforms and teachers and students. When learners use the online education platform to learn, teachers can interact with learners to ask questions, answer questions, etc. Learners can take notes at any time, and learners can also browse each other's notes, reflections, and communication questions. In this way, teachers and students no longer simply complete the class or complete the task after completing the class. Through the interaction between teachers, students and students, the distance between teachers and students is brought closer, the exchange of ideas and the consolidation of knowledge are promoted, which helps to create a good learning atmosphere for students and improve the quality of students' learning.

Compared with traditional classroom learning, online education-based learning methods undoubtedly have great advantages: rich resources, flexible learning time and learning locations, and scientific evaluation. However, it must be said that online learning still has some problems that cannot be ignored, such as the difficulty of ensuring the quality of learning, the failure of students to achieve real learning, the degree of recognition of online learning, and the lack of students' emotions. This requires a learning behavior analysis of these impact factors. The emergence of big data has spawned relevant research on educational big data, which has also promoted research on learning analysis techniques. From the perspective of data, researchers generally consider learner data, learning resource data, and learning behavior data as data that needs to be concerned and studied by learning analysis. Among them, the learner's data generally refers to the learner's personal information, such as educational background, gender, age, place of origin, and so on. The data of learning resources generally refers to the data related to the learning platform and learning content, such as teaching materials, platform functions, teacher strength, and resource update speed. The data of learning behavior is the data generated by the learner during the learning process, such as the amount of clicks, length of study, progress of learning, and activity. As shown in Figure 1:
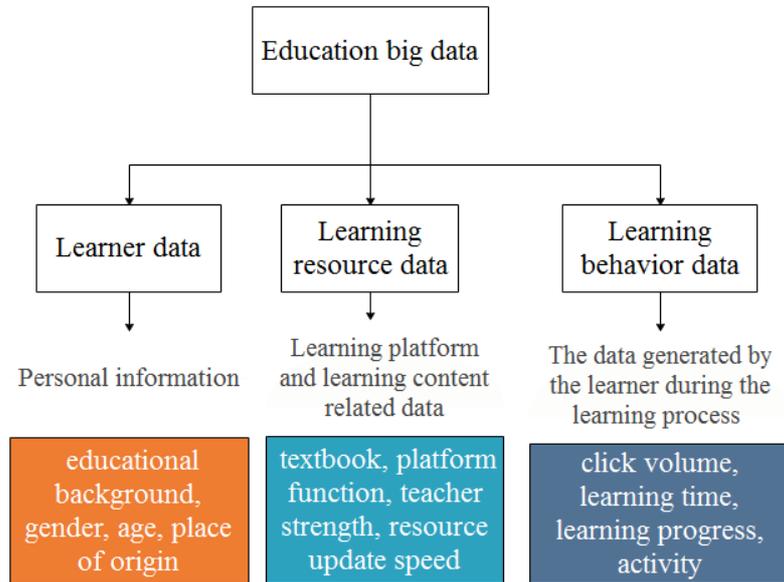
Fig. 2. Research on Learning Analysis Technology in Education Big Data

This paper analyzes the relevant factors that affect the learning behavior of online education students in the big data environment, and uses Pearson correlation coefficient algorithm to calculate the correlation coefficient between online education students' knowledge level and other characteristics of student data, and find the final correlation degree result. Finally, determine the relevant factors that affect the knowledge level of online education students, and propose corresponding control strategies according to relevant factors.

2 Analysis of the Current Situation of Online Education

The survey adopted a random sampling method, and the respondents were mainly college students in Wuhan. Through in-depth campus and on-site questionnaires, 400 questionnaires were distributed and 386 valid questionnaires were returned.

The questionnaire consists of four aspects: one is the basic information of the respondent; the other is the attitude of the respondent to online education; the third is the use and purpose of the online course of the respondent; the fourth is the time and study of the online learning of the respondent Scenario; five is the effectiveness of the online learning of respondents; sixth is the number of online courses that have been completely studied.

Table 1. Questionnaire information statistics

Survey project		Proportion
Basic information of respondents	Male	51%
	Female	49%
Attitude towards online education	Very much in favour	26%
	In favour	32%
	General	21%
	Disapproval	16%
	Very disapproval	5%
Online course usage	frequently used	19%
	Occasionally combined with teaching materials	22%
	Course needs	30%
	Basically not used	17%
	Not used at all	12%
Time for online learning	≥ 3 hours	2%
	2-3 hours	18%
	1-2 hours	23%
	0.5-1 hour	31%
	<0.5 hours	26%
Online learning learning scenario	Before sleep	65%
	When you are in the car	57%
	Lunch break	50%
	After dinner	46%
	Waiting in line	38%
	Wake up	30%
	When going to the toilet	17%
	While eating	9%
	Others	2%
Learning effectiveness of online learning	Helping yourself is to enrich your knowledge and improve your level.	76%
	Improved information search and collection capabilities	69%
	Improved self-learning ability	53%
	Developed critical thinking and problem solving skills	50%
	Developed self-discipline in the network learning environment	42%
	Increased creativity and innovation	31%
	Improve communication and collaboration with people	24%
	Others	3%
Number of online courses that have been fully studied	5 courses or more	6%
	3-5 courses	19%
	1-2 courses	28%
	0 courses	47%

Through the questionnaire, it can be clearly seen that most students are in favor of online learning; more than one-third of the groups using online courses use online courses in order to obtain credits within the prescribed courses. The initiative is not strong and exists for completion. The situation of "forced" use of tasks; online learning time and learning scenarios can be seen that online learning has not become the

main learning method for learners, and is still in a secondary position, but the fragmentation time is valued by learners, more than 1/ 3 respondents can use the fragmentation time to conduct online learning; in the statistics of online learning effectiveness, it can be seen that online learning has a positive impact on learners, and learners are also recognized for the effectiveness of online learning. As can be seen from the number of complete online courses, almost half of the students did not complete a course on the online education platform.

3 Analysis of Factors Affecting Students' Learning Process

The questionnaire made it clear that almost half of the students did not complete a course on the online education platform. The causes of this phenomenon can be analyzed one by one from the emotional conditional factors and the material conditional factors, and the correlation between these factors can be found through analysis. as shown in picture 2:
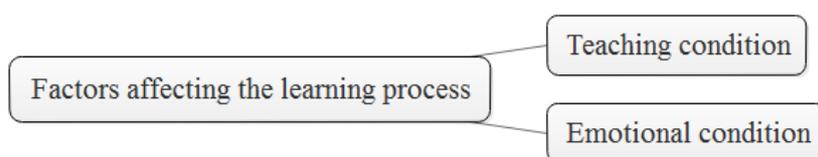
Fig. 3. Factors affecting students' learning process

3.1 Influencing factors of teaching conditions

The influence of teaching conditions on the learning process includes three aspects: teaching method design, functional support and material conditions. From the perspective of teaching, it mainly includes the formulation of teaching objectives, the specific content of teaching, teaching media, teaching methods and teaching activities. From the standpoint of the students, the establishment of teaching objectives should be consistent with the learning objectives of the students, so that the learners can really benefit, so it is the primary factor affecting the learning process.

Generally, the textbooks and course content in the online education platform are generally presented by means of graphics and video, and video is the main one. Through the questionnaire, most students use the fragmentation time to learn, so the video time should meet the objective conditions of most learners. For example, if the video time is too long, on the one hand, the learner cannot complete the complete learning, on the other hand, it will also deplete the learner's patience to a certain extent, so that the learner can't concentrate, and if the video time is too short, it will affect the other side. Teaching progress, on the other hand, will also make learners operate too frequently. The usual course video is usually 30-45 minutes long. From the perspective of the teaching media, the choice and setting of the teaching media can enable the learners to interact effectively with the platform, teachers and other learners, the window of the video play, the location of the course reminder, the en-

trance of the post bar, the entrance of the forum, The window in which the teacher and the learner talk, whether the location of these channels is reasonable. Whether the clarity and size of the video is reasonable and whether there will be a stuttering situation will affect the learner's participation in the course. Whether the arrangement of the teaching process can match the needs of the learner, including the various teaching links, such as the number of assignments, the frequency of quizzes, the frequency of answering questions, whether the teaching activities can fully mobilize and improve the enthusiasm of the learners. Whether the difficulty of the work matches; whether the teaching medium can meet the learning needs of the learner's mobile terminal, and whether the learner needs to communicate through instant messaging software. Whether the teaching method can effectively stimulate the learner's internal learning motivation, and whether the learner's interactivity is rewarded accordingly to improve the viscosity between the learner and the platform. Whether the method of teaching evaluation is diversified and reasonable is whether it can effectively evaluate the learner's learning process and learning outcomes.

3.2 Emotional condition

The emotional condition factors mainly include: social network, including between learners and other learners, between learners and teachers, managers, learners and platforms; learner participation, including learners' attitudes and learning enthusiasm And a sense of accomplishment; whether the arrangement of the learning process can match the needs of the learner, including the various teaching links, such as the number of assignments, the frequency of quizzes, the frequency of Q&A, and whether the teaching activities can be fully mobilized and improved. Whether the learner's enthusiasm and the difficulty of the work match; whether the teaching medium can meet the learning needs of the learner's mobile terminal, and whether the learner needs to communicate through instant messaging software. Whether the teaching method can effectively stimulate the learner's internal learning motivation, and whether the learner's interactivity is rewarded accordingly to improve the viscosity between the learner and the platform. Whether the method of teaching evaluation is diversified, reasonable, and whether it can effectively evaluate the learner's learning process and learning outcomes.

In terms of increased emotional engagement, teachers, managers, and learners need to work together, and good relationships between them help to increase the stickiness of learners and teachers, platforms, and other learners. In many studies at home and abroad, quality education is advocated, and the level of achievement of learners is not ranked and graded. However, in the learning environment of the online education platform, the main purpose of ranking or grading is not to urge learners with slower learning progress, but to encourage students with faster learning and higher enthusiasm through such rankings. To a certain extent, improve the learner's sense of accomplishment, affirm the contribution of learning, make the learner feel a sense of belonging, and create a good competitive environment. The online education platform needs to strengthen the social function of the platform and enhance the interaction between learners and other actors. In the classroom atmosphere, the learning environment of

online education cannot reach the atmosphere of traditional classroom education. The online education platform should strive to create a tradition. A similar atmosphere in classroom teaching, such as group discussions, course forums, and post bars, are there enough learners to discuss, whether teachers and managers can manage and respond to forums and post bars in a timely manner, and whether the questions asked by the learners in class can Get a timely response and get the right time. Whether learners can help and answer each other. In addition to online education, whether the online education platform can organize offline activities in a timely manner to face-to-face teaching to enhance the emotional participation of learners. Generally speaking, the emotional aspect affects the learner's learning process is often the influence of the learner's internal driving force. The positive influence can make the learner invest in the course learning with a positive learning attitude, while the negative influence is It may make learners lose interest in learning, lose confidence, and struggle to solve problems.

3.3 Correlation algorithm design

In the process of differential analysis of different influencing factors, we use the Pearson correlation coefficient algorithm. The Pearson correlation coefficient was proposed by Pearson in 1895. It is also called the Pearson product moment correlation coefficient. In statistics, the Pearson correlation coefficient is widely used to calculate the linear correlation coefficient between two variables, which is defined as:

$$r = \frac{Cov(A_1, A_2)}{S_{A_1} S_{A_2}} = \frac{\frac{1}{n} \sum_{i=1}^{n} \left[\left(x_{i1} - \overline{x}_1 \right) \left(x_{i2} - \overline{x}_2 \right) \right]}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(x_{i1} - \overline{x}_1 \right)^2} \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(x_{i2} - \overline{x}_2 \right)^2}} \tag{1}$$

Where Cov(A1, A2) represents the covariance of variables A1 and A2, X1 and X2 represent the mean of the data in variables A1 and A2, respectively, X1 is defined as shown in equation (2), and X2 is defined as equation (3). Show.

$$\overline{x}_1 = \frac{\sum_{i=1}^{n} x_{i1}}{n} \tag{2}$$

$$\overline{x}_2 = \frac{\sum_{i=1}^{n} x_{i2}}{n} \tag{3}$$

SA1 and SA2 represent the standard deviation of the variables A1 and A2, where SA1 is defined as shown in equation (4) and SA2 is defined as equation (5).

$$S_{A_1} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_{i1} - \overline{x}_1\right)^2}$$

(4)

$$S_{A_2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_{i2} - \overline{x}_2\right)^2}$$

(5)

Using the formula (1), the Pearson correlation coefficient between the variables A1 and A2 can be obtained. The Pearson correlation coefficient is between -1 and 1, that is, the value of r is [-1, 1]. Among them, -1 indicates complete negative correlation, 0 indicates no correlation, and 1 indicates complete correlation. The larger the absolute value, the stronger the linear relationship.

3.4 Significant test

A significant test of the correlation coefficient of the sample can be inferred whether the two variables in the population are related. In practice, because the variance of the overall variable is unknown, the t-test method is often used to test the significance of the sample correlation coefficient. The t value of the t test is defined as shown in equation (6). Where r is the sample correlation coefficient and n is the number of samples.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

(6)

Perform a t-test on the correlation coefficient, first calculate the value of t according to formula (6). Secondly, according to the given significance level a and degree of freedom (n-2), find the corresponding critical value ta/2 (or p value in the t distribution table. If |t|>ta/2 or (p<a), It is shown that r is statistically significant; if |t| ≤ ta/2 or (p ≥ a), it indicates that r is statistically insignificant.

3.5 Correlation result analysis

Through the Pearson correlation coefficient algorithm, the positive correlation of the influencing factors of the online learning process is obtained, as shown in Table 2.

Table 2. Online learning process influencing factors positive correlation

Influencing factor	project	R value
Instructional design	Teaching activities and teaching strategies	0.78
	Evaluation System	0.43
Platform function	Learning support service	0.65
Social factor	Teachers and students	0.83
Learning emotion	Learning information	0.76

4 Strategy Analysis of Learning Process Control Strategy Based on Big Data

According to the correlation analysis, the corresponding control strategy is proposed for specific impact factors. Through the above analysis, it is found that there are several aspects that need to be improved in the evaluation results of the influencing factors based on the learning process control influence factor model. as shown in Table 3.

Table 3. Control strategy analysis

Influencing factor	Project	Control Strategy
Instructional design	Teaching activities and teaching strategies	Need to be rich, and according to the learning style characteristics to push learning materials presented by different organizations or different media
	Evaluation System	Needs improvement
Platform function	Learning support service	Visualize the learner's current learning progress status Visualize the average learning progress of class learners Visualize the gap between learners' progress and average progress Remind slow learners
Social factor	Teachers and students	Encourage and remind learners with low levels of interaction
Learning emotion	Learning information	Visualize learners' academic performance Visualize the average grade of the current class or grade Visualize the relative level of learners in the current class or grade

5 Analysis of Results

The subject of this experimental study was 1862 students in the public elective course of 17th and 18th computer majors in a university in Wuhan, including 982 students in 17th grade as the control group and 880 students in the 18th grade as the experimental group. This comparison experiment ignores individual differences between students of different ages.

The experimental method is that the control group adopts the traditional online education mode, and the experimental group adopts the improved online education mode. The experimental group upgrades and improves the online education system of the public elective course of a university in Wuhan according to the provided control strategy. The improved online education learning situation and the online education in the traditional sense can be compared from the number of logins in the system background. Comparative analysis of learning time, learning performance and academic performance.

5.1 Login times

Table 4. Number of landings in a different chapter of the course (times)

Name	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6
Test group	14	17	16	19	23	15
Control group	14	13	14	15	16	14
Increase	0.00%	30.77%	14.29%	26.67%	43.75%	7.14%

Login number histogram, as shown in Figure 3.
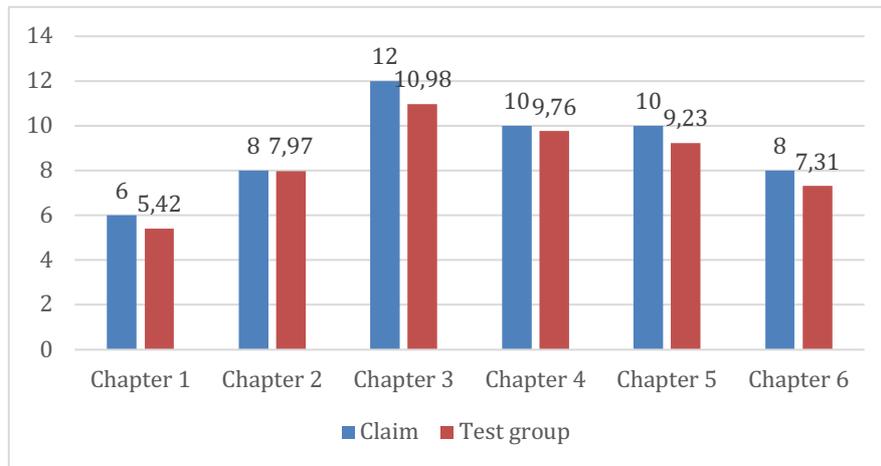
Fig. 4. Login number histogram

The results showed that after the implementation of all control strategies, the number of course-student logins was generally higher than that of the control group.

5.2 Length of study

Table 5. Analysis of the length of study in a different chapter of the course (h)

| Name	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6	Total period
Claim	6	8	12	10	10	8	54
Test group	5.42	7.97	10.98	9.76	9.23	7.31	50.6
Control group	4.21	6.13	10.02	8.31	8.65	7.03	44.35
Increase	28.74%	30.02%	9.58%	17.45%	6.71%	3.98%	14.25%

Login number histogram, as shown in Figure 3.
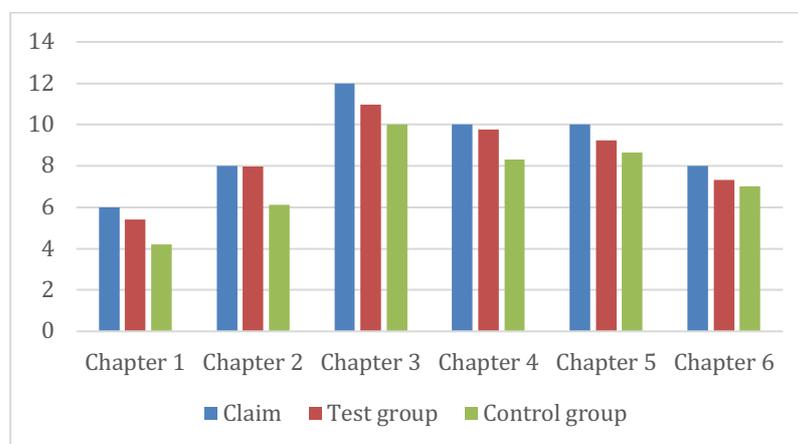
Fig. 5. Login number histogram

The results showed that after the implementation of all control strategies, the course-learner's learning time was generally higher than the control group.

5.3 Academic achievement

Table 6. Analysis of the final assessment results of the first phase of the course

Name	Average score
Test group	83．1
Control group	79．2
Increase	4．9%

The results showed that after implementing all control strategies, the total score of the course-student increased by 3.9 points, higher than the previous average. According to the data in the table, and the evaluation of the course has been improved, the increase is higher than the previous average.

6   Conclusion

Online education is rich in resources and has no time, space, age and other restrictions. Therefore, it has a strong openness. Students can use the online education platform to learn independently, self-monitoring and self-motivation to give full play to students' autonomy. Based on the online education, this paper finds that when students use the online education platform, the effect of learning is not obvious, and there are various problems. Based on these problems, the influencing factors were identified, and the Pearson correlation coefficient algorithm was used to analyze the correlation factors. According to the analysis results, the corresponding learning control strategies are proposed, and the relevant strategies are applied to the online educa-

tion system of a certain course. Finally, taking the online education of computer science in a college in Wuhan as an example, comparing the improved online education system with the traditional education system, it is found that the improved number of logins, learning time and learning interaction have a large increase. It is proved that the improved online education learning system can enhance the sense of on-site teaching, improve the systemicity of learning, improve the completion rate of the curriculum, optimize the quality of education, reduce the pressure on teachers, optimize the certification system, and change the educational concept. Through these improvements, not only can online education break through the current development bottleneck, but also achieve its own rapid development.

7 Acknowledgement

8 References

[1] Hou Changhai. Analysis of China's online education market in 2015. Internet World, 2016(2): 85-88
[2] Huang Wenzhi, Zhao Jing. Discussion on the development prospect of online education in the era of "Internet +"[J]. China Adult Education, 2016(6): 138-140.
[3] Dinah G. & Daniel K. & Isabel W. & Jayne C. Training Needs Analysis: A Literature Review and Reappraisal. International Journal of Nursing Studies, 2014, 42(5): 1-10
[4] Wu Xiaoqiong, Xing Yanfang. Analysis of the Development Path of Online Education[J].Journal of Zhangzhou Teachers College,2015(1):87-90.

[5] Chen Qilin, Bao Haobo. The history and current situation of online education development in China. Xueyuan, 2014(26): 184-185

[6] Yang Wei, Yan Jin. Analysis of the Teaching Reform of Open Online Course——Taking L College as an Example[J]. Value Engineering, 2016, 35(19): 207-209.

[7] Xiao Yuewen, Wang Mingyu. Talking about the status quo and development trend of online education. China Business Theory, 2017(18): 176-177

[8] Earl L., Katz S. Leading Schools in a Data Rich World. Corwin Press, 2006, 23(2): 9-16

[9] Zhang Junchao. Institutional Research and University Management in the Age of Big Data[J]. Higher Education Research, 2014(1) : 128.

[10] Chen Lei. Application of Big Data in Teachers' Online Education Environment——Taking Courses in Zhejiang Province as an Example[J]. Continuing Education, 2017, 31(9): 7-11.

[11] Yang Xianmin, et al. Development Strategy and Path Choice of Wisdom Education in China [J]. Modern Educational Technology, 2014(1) : 14.

[12] Liu Bin, Zhang Wenlan. Research on the Influencing Factors and Structure of Online Course Learning Experience [J]. Modern Educational Technology, 2017, 27(09): 107-113.

[13] Zheng Lei. Online Learning and Quality Assessment Research [J]. Shenzhou, 2014, 14.

[14] Zhang Wei. Strategies for improving the participation of college students in online learning. Zhejiang Normal University. 2011.11-12.

[15] Maclaughlin E J, Supernaw R B, Howard K A. Impact of distance learning using videoconferencing technology on student performance [J]. American Journal of Pharmaceutical Education, 2004, 68(3): 1-6. https://doi.org/10.5688/aj680358

[16] Guan Jia, Li Qitao. Development Status, Trends and Experiences of Online Education in China [J]. China Electro-chemical Education, 2014(08): 62-66.

[17] Sun Li, Cheng Yuxia. Research and Implementation of Online Education Learning Achievement Prediction in Big Data Era--Taking English as an Example of Undergraduate Public Courses[J]. Open Education Research, 2015(3): 74-80.

9 Authors

Zheng Shiyong, comes from Guilin, Guangxi province. PhD student, senior engineer, the main research direction for network marketing, IOT..

Jiang Supring, comes from Hezhou, Guangxi province. Master degree,lecturer, the main research direction for English teaching,Business English,Applied Linguistics in English.

Xiao-Guang Yue is College Advisor of Rattanakosin International College of Creative Entrepreneurship, Rajamangala University of Technology Rattanakosin, Thailand. He is also the IETI Co-Founder.

Ruihui Pu is PhD Candidate of National Institute of Development Administration, Thailand.

Li Biqing comes from Nanning, Guangxi province. Master degree, senior engineer, the main research direction for broadband MINO, IOT,Big data analysis.