

# Review of Online Learning Behavior Analysis

Qianqian Cui

School of Economics and Management  
Beijing Jiaotong University  
Beijing, China

Juanqiong Gou

School of Economics and Management  
Beijing Jiaotong University  
Beijing, China

**Abstract**—Along with the development of Internet technology, online learning platforms have gradually covered universities at home and abroad. The platforms use information technology to overcome the disadvantages of traditional education restricted by time and space, so that learners can achieve flexible interactive and independent learning activities asynchronously. This paper first analyzes the evolution process of online learning platform and the research status of online learning behavior, and then expounds the existing behavioral data analysis technology and the application value of data. Finally, the research status of online learning is summarized, and the research hot spots and development trends in the future are proposed to provide reference for the establishment of personalized intelligent education model.

**Keywords**—online learning platform; online learning behavior; behavioral analysis; behavioral data

## I. INTRODUCTION

As a new learning form, online learning changed the teaching mode of higher education, bringing great convenience for learners. There are endless researches on online learning platforms, which can be divided into two categories. One is about platform functions, and the other is about learning behaviors. This paper first discusses the development process of online learning platform, and then analyzes the characteristics and influencing factors of online learning behavior. In addition, the paper expounds the existing behavioral data analysis technology and the application value of data. Finally, it summarizes the research status of online learning, and puts forward development trends in the future.

## II. ONLINE LEARNING PLATFORM

The online learning platform is comprehensively reflected in the interactive relationship among learning resources, technology platform and people. According to the interactive integration from simple to complex, the evolution process of online learning platform can be divided into four stages with significant characteristics [1].

The first stage is the application stage of single technology, which shows the relationship between technology and technology. In this stage, primary functions are only implemented at a limited level, and typical applications are common learning resource base and simple learning management. Learners can choose the content and time of study independently. However, due to poor basic computer configuration, the function is limited to resource management,

and the content is mainly text and difficult to share widely. Learners at this stage are one-way recipients of learning resources, learning in a passive task-based way.

The second stage is the integration of comprehensive technology, mainly reflecting the relationship between the platform and platform. Typical applications are various general online learning platforms. This stage realizes the basic support of the teaching process: courseware making and publishing, teaching interactive, learning support and teaching evaluation. And it realizes the cross-platform content sharing and integration. However, the platforms in this stage are still relatively independent, so that the connectivity and integration of the systems remains low.

The third stage is the technology disappearing period, which mainly reflects the relationship between platform and people. In this stage, multimedia resources production and release tools, learning analysis and evaluation tools and learning style preference tracking and analysis tools were produced, realizing seamless learning of cross-platform and multi-terminal fusion. Although platforms no longer exist in isolation as they did earlier, they still lack the information architecture support capabilities associated with learning scenarios.

The fourth stage is the emergence of swarm intelligence, which reflects the relationship between people. At this stage, information technology, artificial intelligence and other highly developed. And all learning resources have been digitized and produced. At this stage, people and technology are deeply integrated, and all platforms and network resources achieve seamless integration. Platform and traditional teaching system will no longer be the confrontation between subversion and being overturned, but will jointly shape a new model of human learning and knowledge creation.

## III. ONLINE LEARNING BEHAVIOR

### A. Online Learning Behavior Characteristics

1) *Learning style*: Learning style is the individual characteristic of learners when studying and solving their learning tasks, affecting learners' learning effect. According to the theory of Felder-Silverman learning style, we can analyze learners' learning style from the four dimensions of information processing, information perception, information input and information understanding as follows [2].

a) *Dimension of information processing:* Information processing students have great interest in the information in the online learning platform, whose are sensitive to the opinions expressed by other online learners and the feedback from teachers in the course courseware teaching video within the learning platform. Active learners tend to gain knowledge by actively doing something to discuss or explain ideas to others, and enjoy teamwork, while contemplative learners prefer to learn through deep thinking, either alone or with a regular study partner.

b) *Dimension of information perception:* Students of information perception are accustomed to understand knowledge through subjective perception, and are more accustomed to abstract and interesting learning content. They are greatly interested in video teaching in the learning platform, extended learning articles and communication among students. Perceptive students like to learn facts and pay attention to details, but often choose to avoid complex issues, while intuitive people like to learn theoretical knowledge and have the courage to learn complex things, but they are careless in learning.

c) *Dimension of information input:* Information input students are accustomed to learning by receiving information input from others, and tend to be intuitive and receptive. This kind of students has great interest in reading text or watching video, etc. Visual learners are good at remembering what they see, such as video pictures, etc. Verbal learners, on the other hand, are good at remembering what they hear or read.

d) *Dimension of information understanding:* This kind of students usually process and understand information through their own, which is usually manifested in learning in order or according to their own needs. Orderly learners tend to learn step by step, learning and understanding knowledge in a certain logical order. While comprehensive learners like to think from a global perspective, their thinking is more divergent and jumping.

2) *Learning preferences:* Different type of students for learning has its own preference that affecting students' performance in the online learning platform. Students' preference directly reflects their needs. Primary input learning preferences such as audio-visual and oral learning are relatively easy to meet, and current online learning platforms can basically be implemented. Intermediate preference is mainly the communication activities between individuals and others, including questions from students, answers from teachers and communication between online learners. Advanced preference is the process of independent innovation after students internalize knowledge, such as sharing knowledge independently.

3) *Learning interaction:* In the online learning process, there are two modes of human-computer interaction and human-human interaction. Human-computer interaction behaviors mainly include login, browsing, downloading and other activities that actively acquire platform resources. Person-to-person interaction mainly refers to the statistics of

posting and replying of learners in BBS, which can form a learning interactive network graph. According to the size of learner nodes, the location proximity in the network can be calculated, so as to find out the interaction situation and learning law of the learner and the relationship with other learners [3]. Activity and interaction depth of learners can be used as indicators to evaluate interactive behavior.

#### B. *Factors Influencing Online Learning Behavior*

1) *Learning motivation:* As the main driving force of learning, learning motivation includes two kinds of internal motivation generated by learners' own psychology and external motivation by external forces. Internal motivation is to guide learners to learn independently, while external motivation refers to the reward and punishment mechanism of ranking platform to urge learners to learn through the supervision of teachers. Because learners' self-control ability and environment are different, the learning effect is also very different.

2) *Individual emotion:* Learners' personal emotions also have different influences on learning, which can be divided into positive emotions and negative emotions. Positive emotions can make learners feel happy and satisfied, and thus have great confidence in learning. And negative emotions will make learners feel irritable and generate disgust, thus losing interest in learning.

3) *Learning resources:* Learning resources are the resources that support learning, with various forms of expression and strong interaction. Learning resources occupy a very important position and function, and their rich and diverse choices not only bring convenience to students, but also provide students with more suitable individual learning channels.

4) *Platform performance:* The operational and functional performance of online learning platforms has a great impact on learners' learning behaviors. The platform with less difficulty in operation allows learners to quickly find the learning part they need, saving the search time and improving the efficiency. Good functional performance gives learners more opportunities to choose the most interesting and effective learning methods. A complete academic assistance and support system can help students get correct guidance in the online learning process, reduce the time spent on academic problems, and make students have a better learning experience.

#### IV. ONLINE LEARNING BEHAVIOR DATA

Online learning behavior data is the data generated by the interaction between learners and the platform in the learning process, which is mainly stored by the platform database and other tools in real time. By analyzing the online learning behavior data, teaching stakeholders can have a more comprehensive understanding of the learning process, and realize the scientific prediction, evaluation and intervention of the learning process.

### A. Data Acquisition

The acquisition of behavioral data is the premise of learning behavior analysis. According to the differences of data acquisition objects and policies, it can be divided into server-side and client-side approaches, while from different perspectives of terminals, data sources are divided into mobile terminal, PC terminal and client terminal. Multi-terminal and all-aspect data acquisition methods can help us understand learners' learning characteristics better. The online learning behavioral data acquisition framework is shown in "Fig. 1" [4].

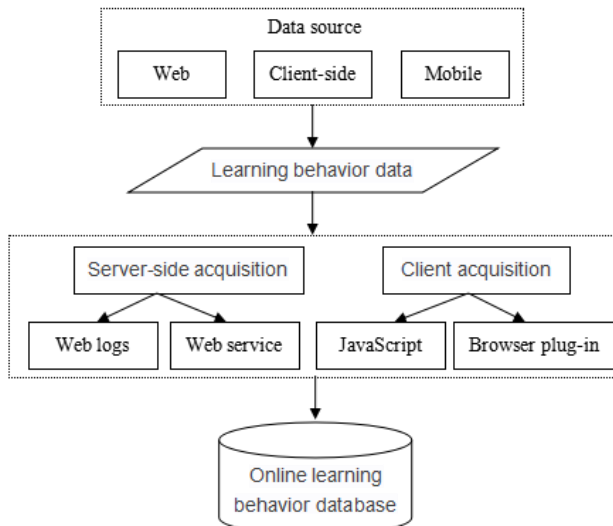


Fig. 1. Online learning behavior data acquisition framework.

Server-side data collection mainly includes Web logging and Web service. Web log is used to store the data of learner's real-time operation, including the user's request time, request type, request content, request status, client's access address, the time of operation execution, the browser type used by the client, etc. Web service is a way to implement data collection through back-end code. Researchers can design the platform module database module according to the learning behavior type, so that the target content can be collected according to the needs, and the acquisition of learner behavior data can be more comprehensive and flexible, with a wide range of services.

Client-based data acquisition mainly refers to the acquisition of data generated by learners when they use the browser to learn locally, which mainly uses Java script Cookie to collect data to get information about the learner's browsing behavior easily. This method stores the behavior data of learners in a specific location and reads the data from the stored information when needed, so as to achieve flexible data collection and record the access of cache proxy server, and to track the behavior of visitors more accurately.

### B. Data Analysis Method

Data analysis methods mainly include educational data mining, statistical analysis, social network analysis and discourse analysis. S.B.S hum et al. used text mining to identify exploratory conversations in synchronous text chat by analyzing text data in the learning process [5]. Park et al. used the clustering method in data mining to carry out the clustering

analysis of behavioral data to extract the behavioral characteristics of learners [6]. Zong Yang et al. found the relationship between learning behavior and learning effect by using logistic regression method, and obtained the indicators to predict learning effect [7]. Shi Yuefeng et al. analyzed group learning interaction from the perspective of social network and found interaction rules to improve learning quality [8]. Zuo Mingzhang et al. constructed an interactive discourse analysis model by means of discourse analysis method to intervene in learning behavior [9].

1) *Data mining*: Data mining conducts research on large-scale disordered data in the database, so as to discover the hidden and valuable deep contents in the data. The main function of data mining is to describe and predict. Description is to show the characteristics of data, while prediction is to predict the value of predicted data [10]. The basic steps of data mining are shown in "Fig. 2".

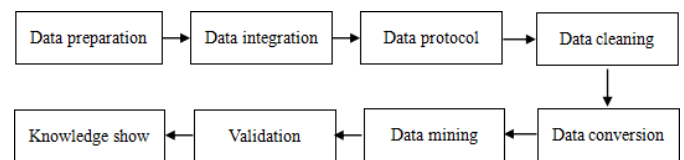


Fig. 2. Basic steps of data mining.

2) *Statistical analysis*: This method refers to the collection of a series of logical thinking methods and quantitative relationship analysis methods. Behavioral data analysis mainly uses descriptive statistical analysis correlation analysis regression analysis to discover the relationship between learning data. Descriptive statistical analysis can describe the basic characteristics of learning behavior and discover the basic state and law. Correlation analysis finds out whether there is a dependency relationship between variables through the correlation analysis of behavioral characteristics learning behavior and learning effect. Regression analysis is the regression analysis of learning behavior and learning effect to study the causal relationship between behavior and effect, then analyze which behavioral indicators can predict the learning effect, and establish a predictive mathematical model.

3) *Social network analysis*: Social network analysis is to analyze the interaction behavior of online learning behavior. In the online learning platform, the interactive behavior of learners takes place in BBS interaction, which is difficult to fully express the interactive relationship between teachers and students through the original data. The degree of correlation between teachers and students and the impact of individual behaviors on other individuals need to be analyzed through interactive data [11]. The analysis process is shown in "Fig. 3".

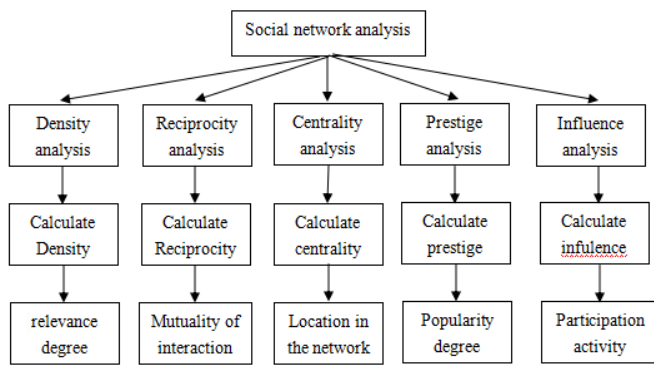


Fig. 3. Social network analysis process.

4) *Discourse analysis*: As the most direct communication carrier between learners and the explicit manifestation of inner psychological processing, discourse is the main data source for studying interactive learning behaviors. Discourse analysis can deeply explore the rules of learners' interactive discourse behavior in BBS. It adopts content analysis method to analyze the posting behaviors of learners in BBS, and classifies them. Then, through probabilistic topic modeling and word frequency analysis, it extracts hot learning topics in the interactive discourse process of learning groups, and compares the amount of posts and emotional experience (the difference between positive and negative emotions) of learners with different learning outcomes in different time periods [12]. Finally, there is a correlation between behavioral variables and learners' learning outcomes, so as to guide online learning behaviors.

### C. Application Value of Data

The behavioral data generated by online learning can truly reflect learners' learning situation, learning progress, learning rules and learning effects. The analysis results have utilization value for both learners and platform builders. Currently, the application and value of data are mainly reflected in the following four aspects.

1) *Assessment and prediction of academic performance*: As an important link in education, performance evaluation is also an important part of online learning platform. Existing performance evaluation methods only pay attention to the results evaluation and ignored the process evaluation, there are a lot of bias. Therefore, based on the data of online learning behavior and text data, determining the effective learning achievement evaluation index to evaluate learners' learning and prediction, will help students and teachers to understand students' learning situation, and the evaluation of the course final grade to provide beneficial reference value.

2) *Learning behavior guidance*: Through the analysis of learning behavior in the process of data, we can find learners' intrinsic characteristics of the rules of learning and the intrinsic learning characteristics of learners. We can also present the analysis results to teaching stakeholders through visualization technology, so that learners can clearly understand the interaction between self-learning and

classmates. If there is any deviation, timely adjust the learning progress and learning style. For researchers, by analyzing the behavioral data of the online learning of graduate students, they can clarify their research direction and progress. When the graduate students browse, upload or comment on the platform deviating from the research topic recently, they can give timely reminders to standardize and guide the research and learning behaviors.

3) *Teaching content improvement and improvement*: According to the online learning behavior data analysis results, the teachers can understand learners' learning preference, learning characteristics, learning styles and so on. At the same time, the teachers also can know students mastery of curriculum knowledge. According to these, the teachers can adjust the teaching method and content at any time, explain the key points that are not easy to master in the analysis results, and quickly pass the points that are easy to be understood by most learners. This can not only save teachers' time in preparing for the key analysis, but also bring better teaching results.

4) *Online learning platform function improvement*: There is a co-evolutionary relationship between online learning platform and learner behavior. The function of the platform has a great influence on the learner's learning behavior, and the analysis of the learner's behavior can provide reference value for the improvement of the platform function. For example, the analysis of online learning behavior data can be used for performance evaluation and learning effect prediction, thus improving the performance evaluation function of the learning platform. And the analysis of learning behavior characteristics can provide data basis for the platform's personalized intelligent learning content push learning guidance and other functions. With the enrichment of online learning resources, online learning behaviors have gradually evolved. Online learning platforms have gradually absorbed such tools as email blog, BBS social network, and gradually become an open learning management system.

## V. CONCLUSION

To sum up, with the development of information technology, the function of online learning platform is gradually improved to promote the process of online learning and affect the learning effect of learners. In the process of learning based on the platform, learners generate a large amount of behavioral data. The specification of learning resources and data enables the platform to store learning behavioral data in a more comprehensive way, so as to more accurately analyze the characteristics of learners' behaviors. Then we can predict the learning effect and evaluate the learning process to provide guidance strategies for learning intervention and personalized learning.

In recent years, with the rapid development of artificial intelligence theory and computing power, artificial intelligence application combined with cognitive science has become an important research and application direction of online learning platform in the future. Machine learning and intelligent

optimization algorithms will play an important role in the teaching process of large-scale education. Many foreign universities represented by Stanford University have tried and studied in this field. On the one hand, they analyze and draw knowledge maps through machine learning technology. On the other hand, through artificial intelligence and big data application technology, it identifies and analyzes the learning cognitive process patterns of learners, providing more effective support for learners and managers of institutions. At home, according to the research results of the online education research center of the ministry of education, THU has also launched the artificial intelligence teaching assistant small wooden robot and real-time intelligent classroom teaching interactive platform rain classroom. Based on such applications as rain classroom, intelligent online learning platform oriented to the future is built through the platform of teaching process and artificial intelligence technology. This requires that future-oriented online learning services should take open platform as the concept and excellent content as the core carrier, realizing the integration of the whole education and teaching process through technical means. The platforms in the future will support the application and implementation of a variety of mixed teaching methods so as to truly realize the scale of personalized education and high quality.

#### REFERENCES

- [1] Han Mo. Research on online learning platform [D]. Hebei Normal University, 2019.
- [2] Yuan Shuai. Online learning behavior analysis and evaluation and its application research [D]. Central China Normal University, 2011.
- [3] Liu Bing, Li Yanming. Analysis of SPOC online learning behavior characteristics of learners — Based on interactive perspective [J]. Journal of Jimei university, 2019, 20(01): 33-38.
- [4] Wang Hui. Online learning behavior analysis and applied research [D]. Central China Normal University , 2018.
- [5] Shum S.B. , Ferguson R. Social Learning Analytics [J]. Journal of Educational Technology&Society, 2012, 15(3): 3-26.
- [6] Park Y, Yu J H, Jo I. Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute [J]. The Internet and Higher Education, 2016, 29.
- [7] Zong Yang, Chen Li and Zheng Qinhu. The study of distance learner's academic emotion analysis based on online learning behavior data — Take Moodle platform [J]. Open learning research, 2017, 22(06): 11-20.
- [8] Shi Yuefeng, Liu San and Liu Zhi. An empirical study on online learning behavior analysis based on social network analysi [J]. China Education Information Network, 2019(01): 5-10.
- [9] Zuo Mingzhang, Zhao Rong and Wang Zhifeng. Construction and practice of interactive discourse analysis model based on BBS text [J]. E-education Research, 2018, 39(09): 51-58.
- [10] Fu Gangshan, Wang Gaihua. Research on network learning behavior and learning effect based on data mining [J]. E-education Research, 2014,35(09): 53-57.
- [11] Liu San, Shi Yuefeng and Liu Zhi. Application research of group interactive learning analysis in network environment — based on the perspective of social network analysis [J]. China Educational Technology, 2017(02): 5-12.
- [12] Liu Zhi, Zhang Wenjing and Sun Jianwen. Study on learner interactive discourse behavior analysis in cloud classroom BBS [J]. E-education Research, 2016, 37(09): 95-102.