A REVIEW OF BEHAVIOR ANALYSIS OF COLLEGE STUDENTS

Wei-hong WANG, Hong-yan LV, Yu-hui CAO, Lei SUN and Qian FENG

School of Information Technology, Hebei University of Economics and Business, Shijiazhuang, China

ABSTRACT

Student behavior analysis plays an increasingly important role in education data mining research, but it lacks systematic analysis and summary. Based on reading a large amount of literature, this paper has carried out the overall framework, methods and applications of its research. Comprehensive combing and elaboration. Firstly, statistical analysis and knowledge map analysis of the relevant literature on student behavior analysis in the CNKI database are carried out, and then the research trends and research hot spots are obtained. Then, from the different perspectives of the overall process and technical support of student behavior analysis, the overall framework of the research is constructed, and the student behavior evaluation indicators, student portraits and used tools and methods are highlighted. Finally, it summarizes the principal applications of student behavior analysis and points out the future research direction.

KEYWORDS

Student Behavior; Knowledge Graph; Behavior Analysis; Student Portraits; Data Mining.

1. INTRODUCTION

Education is the foundation for a hundred years. Both the state and society attach great importance to education. General secretary Jinping Xi mentioned education 43 times in the report of the 19th national congress. Students occupy a dominant position in the process of education and teaching, and their behavior performance can reflect the level and quality of education, so the analysis of student behavior is particularly important. However, in the traditional teaching model, it is difficult to record and collect the relevant behavioral data of students. With the continuous development of information technology, artificial intelligence as the big data and cloud computing and other technology to promote the innovation and development of all walks of life, as early as the end of the twentieth century in education field gradually to the informatization development, in recent years, the state has staged a series of about the development of education informatization policies, which are the key mentioned vigorously develops the education informationization, realize the fusion of emerging information technology and education, promote the education teaching reform, improve teaching quality of education. General secretary Jinping Xi has also pointed out that China should make unremitting efforts to promote the informatization of education and expand the coverage of high-quality education resources by means of informatization.

After more than 20 years of development, digital education resources have become increasingly abundant, educational management informatization has achieved remarkable results, and the
application of educational informatization has been continuously deepened [1]. In the process of
education and teaching, a large number of students' behavior data have been accumulated.
Nowadays, with the continuous development of information technology, the analysis of these
data by using modern information technology can provide a more comprehensive understanding
of students and timely find out the deficiencies in teaching, which is conducive to the
development of students and the reform of teaching mode. Therefore, there are more and more
studies on student behavior analysis, and a large number of research results have been
accumulated so far. However, it can be seen from the existing literature that although there are
many research results on the behavior analysis of college students, there is still a lack of in-depth
combing and systematic summary of the existing research. Therefore, it is the original intention
and foothold of this paper that how to make full use of the existing achievements to provide
effective guidance for engineering practice and how to let the later researchers have a
comprehensive understanding of the overall framework and shortcomings of the analysis of
college students' behavior in order to better carry out the follow-up research.

In this paper, the main content: part 2 research on behavior analysis in the CNKI database
statistics of the number of relevant literature and use CiteSpace V tool knowledge map analysis,
and concluded the research tendency and research hotspot; In the third part, from the perspective
of the general process and technical support of student behavior analysis and research, the overall
frame diagram of the study is constructed by reading and analyzing relevant literature, and the
indicators of student behavior evaluation, student portraits, and the tools and methods used are
mainly introduced. Part 4 summarizes the important application of student behavior analysis and
points out the future research direction. Part 5 summarizes the whole paper.

2. LITERATURE ANALYSIS OF STUDENT BEHAVIOR ANALYSIS RESEARCH

2.1. Statistical Analysis of Relevant Literature on Student Behavior Analysis

The topic of "student behavior analysis" was searched in CNKI database, and 477 articles of
pertinent literature on student behavior analysis were obtained. On this basis, statistical analysis
was conducted. The trend statistics of the topic of "student behavior analysis" were shown in
figure 1. It can be seen from the figure that as early as the 1970s and 1980s, some scholars have
paid attention to student behavior analysis, and since the 21st century, papers with the theme of
"student behavior analysis" have shown an increasing tendency.

![Figure 1. Statistics of the trend of the topic "student behavior analysis"](image-url)
2.2. Knowledge Graph Analysis of Students' Behavior Analysis Research

CiteSpace is a bibliometrics tool that can be used to observe research trends or trends in a field and present them visually. In this paper, the relevant literature on "student behavior analysis" from 2008 to now in CNKI database is used as the data source. After excluding some has nothing to do with the education field of literature, the remaining 387 articles as input data, with the aid of CiteSpace V tools, a clustering analysis of the literature can be seen from the key words in the knowledge map in the study of student behavior analysis of relevant keyword occurrences more is, student behavior, behavior analysis, data mining, big data, etc., followed by association rules, K - Means algorithm, campus network user behavior, performance prediction, etc. It can be seen that the research of student behavior analysis mainly USES data mining tools or algorithms (such as association rule algorithm, clustering algorithm, etc.) to analyze students' learning data, campus network data and other teaching data in colleges and universities so as to predict students' relevant behaviors.

3. RESEARCH FRAMEWORK OF STUDENT BEHAVIOR ANALYSIS

3.1. The Overall Framework of Student Behavior Analysis

Student behavior analysis is tantamount to model and analyze a large number of student behavior data to get the feedback of students' behavior correlation or students' learning interest and learning effect to relevant teachers and students and university administrators. Student behavior analysis is an application of data mining in a university environment. Through reading and analyzing a large number of literature about students' behavior analysis [4,5,18-22] summary analysis found that most students behavior analysis of student behavior data collection work first, and then to the original data preprocessing and feature selection, and according to different research goal to build the relevant student behavior index, then clustering, correlation analysis and other related algorithm to establish related model and validation, finally apply it to education of colleges and universities teaching actual scenario. The overall framework of student behavior analysis is illustrated in Figure 2.

![Figure 2. The overall framework of student behavior analysis](image-url)
3.2. Constructions of Student Behavior Evaluation Index

3.2.1. Indicators of Student Behavior Evaluation

In the analysis of students' behavior data, different index systems should be constructed according to different analytical objectives. In the construction process of students' behavior evaluation index, relevant knowledge and experience in pedagogy, psychology and other fields will be engaged. In addition, Delphi method, expert ranking method, principal component analysis method, analytic hierarchy process and other methods will be utilized. In the research of student behavior analysis, many scholars have been conducted relevant research on the construction of student behavior evaluation index [2–5].

3.2.2. Construction Process of Student Behavior Evaluation Index

By reading a large number of relevant literature, the student behavior evaluation index construction process as shown in Figure 3, through the pedagogy, psychology and other fields related knowledge, experience and related literature to determine the preliminary evaluation index, and then by using the methods of principal component analysis (revised preliminary determine the evaluation index and determine the weight of each index, it is concluded that the final student behavior evaluation index.

![Figure 3. Construction of student behavior evaluation index process](image)

3.3. Student Portrait

The student portrait is a continuation of the user portrait in the application of big data in colleges and universities. According to the data of students' behaviors in school, a tagged student model is abstracted. With the update of student behavior data, student portrait is also changing dynamically. Student portrait lays a foundation for student behavior analysis. Especially in the era of massive data, the construction of student portrait is very meaningful for the education and management of students [6]. Student portraits can be divided into individual portraits and group portraits (as showed in Figure 4). Unique portrait of students is the modeling of a certain student, while group portrait of students is the modeling of a certain class, grade or major.

![Figure 4. Individual portraits of students and group portraits of students](image)
3.3.1. Individual Portraits of Students

As early as the end of the 20th century, some scholars conducted research on student behavior analysis through student portrait [7]. With the development of information technology, there are more and more techniques to construct student portrait and more and more contents of student portrait. The individual student portrait actually labels a student based on student behavior data. Through the individual portrait of students can better understand the personalized development of students, improve the teaching level, and make education really into teaching according to aptitude. The researchers mainly use the data of students' basic information, attendance information, Internet access information, book borrowing information and so on to realize the individual portrait of students through processing modeling to serve the teaching activities [8~11].

3.3.2. Students Group Portrait

Individual portrait helps to fully understand every detail of individuals, but is not conducive to the management of the overall grasp. Therefore, group portraits are also concerned. By crunching massive amounts of data, the researchers analyzed the characteristics of student groups and constructed portraits of them [12~13].

3.4. Tools, Methods and Comparisons of Students' Behavior Analysis

3.4.1. Main Tools Used in Student Behavior Analysis

At present, there are many learning analysis tools, such as weka, SPSS, Google Analytics and Mixpanel, which can be classified according to different rules [17]. Many researchers at home and abroad use weka and SPSS to analyze students' behavior [10][14~16]. The use of relevant tools will make the research work twice as effective at half the time. Moreover, the research of students' behavior analysis with the help of tools is not difficult to operate and easy to get started, without too many technical requirements for researchers. However, there are a few limitations. It is impossible to adjust and improve the algorithm according to the unique characteristics of students' behavior data, which leads to the low accuracy of model prediction. Therefore, some scholars will directly utilize some data mining algorithms and their corresponding improved algorithms to conduct research on student behavior analysis.

3.4.2. Data Mining Algorithm in Student Behavior Analysis

Data mining algorithms mainly include clustering algorithm, classification algorithm and association rule algorithm. Different algorithms have their own advantages and disadvantages in dealing with different student behavior data, and all kinds of algorithms have been applied in the study of student behavior analysis, some of which are shown in Table 1.

<table>
<thead>
<tr>
<th>Types of algorithms</th>
<th>Examples of application</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering algorithm</td>
<td>Ding D, Li J et al. analyzed the advantages and disadvantages of k-means clustering algorithm in campus data, optimized it, and designed a clustering method based on density division to prove the effectiveness and practicability of the improved algorithm in practical scenarios.</td>
<td>[18]</td>
</tr>
<tr>
<td></td>
<td>Wang Fayu and Jiang Yan combined the self-organizing neural network with the fuzzy c-means clustering algorithm to avoid the</td>
<td>[19]</td>
</tr>
</tbody>
</table>
error caused by the improper initialization of the fuzzy c-means clustering algorithm, and analyzed the students' learning interest through the campus wireless network data on the Hadoop platform.

Classificati\textunderscore on algorithm
Hu Y H, Lo C L et al. found that the classification and regression tree (CART) supplemented by AdaBoost was the best classifier for the evaluation of learning performance in the study in the experiment of predicting students' online learning performance.

Association rule algorithm
Guo Peng and Cai Cheng used the improved k-means clustering algorithm and the Apriori algorithm which introduced the degree of interest to analyze and mine the student achievement information and the course information.

With the continuous development of education informationization, the student behavior data clearly increased, also more and more complex data types, lead to a simple data mining algorithm in the study of student behavior analysis of low efficiency, accuracy, can't meet the needs of the existing research, represented by artificial intelligence of big data, knowledge map the emergence of new technologies such as solved this problem to some extent, promote the transformation and innovation in the field of education.

3.4.3. Emerging Technologies Such as Big Data and Knowledge Mapping in Student Behavior Analysis

Different technologies have their own advantages. For example, big data platform can realize parallel operation and reduce the time of processing massive data. Knowledge map is conducive to knowledge mining and reasoning. Therefore, many scholars use these new technologies to conduct research on student behavior analysis, as shown in Table 2.

Table 2 advantages and disadvantages of emerging technologies used in student behavior analysis and their application examples.

<table>
<thead>
<tr>
<th>Technical name</th>
<th>advantage</th>
<th>disadvantage</th>
<th>Examples of application</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big data</td>
<td>Parallel computing, online analysis, high real-time.</td>
<td>Configuration requirements are high.</td>
<td>Ding D, Li J et al. proposed the parallel design of their student behavior algorithm by using Spark framework, which improved the efficiency of the algorithm.</td>
<td>[18]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Wang Fayu and Jiang Yan ran the clustering algorithm in parallel on the Hadoop platform when analyzing the students' learning interest through the wireless data of the campus, which effectively reduced the running time of the algorithm.</td>
<td>[19]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cantabella M, martinez-espana R et al. used Apriori algorithm in Hadoop MapReduce framework to conduct student behavior analysis of data in learning management system.</td>
<td>[23]</td>
</tr>
</tbody>
</table>
Knowledge map

Knowledge can be extracted from students' behavior data, and knowledge reasoning and knowledge mining can be carried out.

Need the support of a large number of data, high technical requirements, need the support of domain knowledge.

Su Yu et al. studied the prediction of scores in the assessment of students' academic ability, and took the prior understanding of anchor graph as the regularization item of the self-encoder to provide teaching and research basis for personalized learning recommendation. [32]

Zhai Yu, Xu Meng et al. constructed a knowledge map about the relevance of knowledge points, and used the knowledge map to deduce the knowledge state of learners so as to recommend learning resources. [33]

Different tools and methods have their own advantages and disadvantages. In practical research and application, a variety of techniques and methods should be combined with the actual situation to conduct research and analysis, so as to improve real-time and accuracy.

4. Practical Application of Student Behavior Analysis and Future Research Trend

4.1. Practical Application of Student Behavior Analysis

The practical application of students' behavior analysis includes students' course performance prediction, warning of dangerous behaviors, recommendation of learning resources, etc. At present, many domestic and foreign universities have applied the relevant research results of student behavior analysis to the actual teaching and education environment. Among foreign universities, Marist College in the United States has partnered with Pentaho, a business data analysis firm, to improve graduation rates by collecting and analyzing students' study habits to predict their academic performance and intervene to help troubled students. Admissions officers at Wichita State University in Kansas used big data analysis to forecast the percentage of incoming undergraduates who succeed or drop out. The University of Luxembourg in Germany has improved its research capacity of information education by using the development model of Internet information schooling. In domestic colleges and universities, university of electronic science and technology research and development of students "portrait" system has good use in the campus management, the system can be predicted according to the student behavior students may fail, assist teachers to understand students, can also according to student's consumption behavior that students of poverty, help the school more accurate support work, etc. Xi 'an Jiaotong university also USES a similar system to analyse students' behavioral portraits for performance prediction, accurate funding and behavioral early warning. In addition, Peking University, Beijing normal university, Wuhan university, Shaanxi normal university and other universities have similar applications.

4.1.1. Performance Prediction

Grades are one of the important indicators of students' learning effect, so grades prediction is an important application of students' behavior analysis.
(1) performance prediction based on online learning behavior

In recent years, online learning platforms have developed rapidly. Moocs and other platforms have been well known, and many universities have also introduced online classroom teaching mode. Therefore, many students' learning behavior data are kept in these platforms. In the field of education and academia, there are many researches on the prediction of performance based on online learning behavior. Balakrishnan G and Coetzee D [34] used the hidden markov model to predict student retention of moocs. Kizilcec R F and Piech C et al. [35] also analysed the learner data in the massive open online courses to obtain the learner subgroups.

(2) performance prediction based on offline students' behaviors

With the information and digital management of various departments in colleges and universities, the offline behavior data of students are also recorded, such as the relevant data in the one-card and the campus wireless usage data, so there are more and more researches on the performance prediction based on the offline student behavior. Liu Bopeng, Fan Tiecheng et al. [22] analyzed and modeled data from three aspects, including students' behavior, personal attributes and historical performance, and used support vector machines to give early warning of students' performance. He chu, Song jian et al. [24] analyzed the relevance of courses and predicted student performance. Miao C, Zhu X et al. [36] analyzed students' behavior characteristics on the basis of campus wifi data to analyze their performance.

4.1.2. Behavioral Early Warning

Through the behavior of students can reflect the psychological status of students, etc., in-depth exploration of students' behavior and psychological problems for early warning of students' behavior can timely find the abnormal behavior of students, and assist the production safety control of teaching. Under the Hadoop platform, the early warning platform system of students' campus behavior analysis or early warning and decision system of smart campus can be realized [25][26].

4.1.3. Recommendation of Learning Resources

Through student behavior analysis, we can better understand students' learning methods, learning habits and learning interests, so that we can recommend students' learning resources, mainly online course resources [27–29].

4.1.4. Analysis of Students' Comprehensive Quality and Ability

Hidebound patterns of traditional student evaluation system exist, such as formalism serious shortcomings, with student behavior analysis to guide analysis of students' comprehensive quality helps to establish a scientific evaluation concept, follow the college students' comprehensive quality evaluation of comprehensiveness, the principle of feasibility, etc[30], is advantageous to the student's career planning, but also conducive to the education teaching method improvement. He Yi [3147] constructed the comprehensive quality evaluation system for college students by combining various comprehensive evaluation methods, designed and implemented the evaluation system, and provided a good idea for the comprehensive quality evaluation of college students and the standardized management of student work.
4.2. Future Research Trends of Student Behavior Analysis Research

Through the analysis and combing of the existing literature, it can be seen that the research on student behavior analysis has made some progress, but there are still some deficiencies in the aspects of student behavior evaluation index, related algorithm research, and source of student behavior data set. This is the trend of future research.

(1) The data sources of student behavior analysis are relatively limited, and there is a lack of representative public data sets for scholars to study and analyse.
(2) There is a lack of unified indicators of student behavior evaluation, so researchers need to construct the indicators of student behavior evaluation based on their own analysis objectives and relevant methods.
(3) In related studies, there are more studies on students' personal portraits, which can help to accurately locate students, but there are few studies on students' group portraits of classes, schools and departments, which is not conducive for managers to grasp the overall situation of students. Subsequent scholars can carry out research work in the aspect of students' group portraits.
(4) The relevant algorithm used in the study is relatively simple, and the algorithm research in the field of student behavior analysis is not in-depth enough. The original algorithm can be reasonably improved according to the data to improve the accuracy of the analysis.
(5) With the continuous development of emerging technologies, knowledge mapping has become a hot research issue in all walks of life, but the combination of knowledge mapping and student behavior analysis is less, and the relevant domain knowledge mapping is also less.

5. CONCLUSIONS

The analysis of students' behavior plays an increasingly important role in the research of college education and teaching, but there is a lack of systematic analysis and summary of it, which is not conducive to the relevant researchers to quickly form a knowledge system and conduct in-depth research. On the basis of reading a large number of literatures, this paper makes a comprehensive review of the research results of student behavior analysis in recent years, and presents the research trends and research hotspots of student behavior analysis by means of statistical analysis and knowledge map analysis. Based on the general process, technical support and other aspects of student behavior analysis research, this paper constructs the overall frame of the research, and mainly introduces indicators of student behavior evaluation, student portrait and the tools and methods used. Finally, it summarizes the important application of student behavior analysis and points out the future research direction, which can provide reference and help for researchers.

REFERENCES


© 2020 By AIRCC Publishing Corporation. This article is published under the Creative Commons Attribution (CC BY) license.