Research on Data Mining of Intelligent Education Evaluation from the Perspective of the Knowledge Graph

Yanan Chen
Central South University of Forestry and Technology, software engineering, changsha, 202102, China
E-mail: 1617010665@qq.com

Abstract: The continuous development of Internet information technology given birth to data mining technology. Data mining is applied to analyze the internal relationships of data and provides a more optimized advancement way. The paper focuses on how to improve the data mining and analysis capabilities of intelligent education to associate students of the same characteristics. A cluster analysis method was utilized to construct a students' knowledge graph model, which used a text classification algorithm to classify the evaluation questions and attribute them into various knowledge points. On the basis of the analysis on the data generated from online education, the clustering algorithm was adopted to classify the learning characteristics of the students. A knowledge graph developed by associating the students with similar characteristics was visualized be means of information technology in order to realize visualized and intelligent teaching.

1. Introduction
Intelligent education can also be called smart education, whose common manifestation form is online teaching.

Online teaching makes it possible to mine educational big data, based on which the students’ problems in learning and their understanding degree of knowledge can be fully grasped. In addition, it helps teachers optimize teaching method, fulfill the actual demand of students as much as possible and improve the quality of teaching.

Currently, with the deeper and more extensive application of information science and technology in education, intelligent education has been widely known and researched by educators, but some mistaken concepts exist in a lot of studies. Some people think that intelligent education is to establish a learning platform for students studying online through computers. It is a misunderstanding of intelligent education. In other words, this understanding is not clear, in-depth and comprehensive enough. This research is aimed at realizing intelligent education by gripping the key issues of intelligent education and adopting methods such as data mining, thereby enhancing students’ interest in learning and teachers’ teaching quality. The direction of intelligent education is correct, and it should be an important direction for future education reform. Intelligent education lends more interests to study and heighten teaching efficiency through a variety of methods, which of great significance to the growth of talents and economic and social development.
2. Overview of Intelligent Education Research
Educational intelligence can be understood as comprehensively and deeply applying modern information technology into the field of education to boost educational reform and development. As a transformation of the traditional education model, intelligent education can optimize the quality of teaching by constructing an intelligent model. It can be divided into intelligent learning, teaching and management. Advancing intelligent education should be based on building an intelligent education model, whose main elements include intelligent teaching, learning, management and training. Undoubtedly, only by applying information science and technology can intelligent education be realized, and it is necessary to mine and analyze the data generated from the intelligent education process. Relying on computer information technology, intelligent education adopts specific algorithms for analyzing relevant data to swiftly and accurately dig out the characteristics and problems in educational activities between students and teachers, thus proposing personalized guidance and suggestions for students, which highlights the pertinence and effectiveness of the guidance and suggestions. Therefore, it has become a top priority to realize intelligent education by accelerating the establishment of an adaptive learning system. The adaptive learning system is a concrete manifestation pattern of educational data mining applications, and it is worthy of in-depth research due to its allowing students to make adaptive adjustments according to their own characteristics.

3. Construction of Knowledge Graph Model

3.1. Model construction ideas
The model in this paper was constructed by adopting methods such as text classification and cluster analysis to conduct text analysis on the relationship between test questions and knowledge. Then, based on the data generated by students in a test, their understanding degrees of the knowledge points were transformed into values to form corresponding learning characteristics. The first was to analyze the students’ completion of the test questions, and the second was to study the affiliation between the test questions and knowledge points. Afterwards, the student's test results were integrated to summarize the learning characteristics of them. Students of the same characteristics were classified to construct a knowledge graph of students based on the clustering analysis algorithm.

3.2. Steps of constructing the knowledge graph model

3.2.1. Classification of the test text
The automatic association of test questions and knowledge points can be realized by text classification. There are a large number of questions for a test, but the past manual classification had a quite low speed and poor effect. After the application of text classification, the costs have been reduced largely, and the efficiency is also much higher than manual classification. In this section, the text classification of test questions is composed of the extraction of the mapping relationship between TF-IDF and VSM-based questions and knowledge points, the filtering of participles and stop words, text feature extraction, model classification, result evaluation and classification result output.

3.2.2. Extraction of students’ characteristics
The data generated by students during answering the questions were set as their understanding degrees of the knowledge related to the questions, and a text classification algorithm was used to transform these degrees into the characteristics of students in learning certain knowledge point. After the extraction of these characteristics, the clustering algorithm was adopted to group the students of the same characteristics together for analyzing each type and researching their uniqueness. In this section, relevant elements in the process of researching the test include the rate of correct answers, the length of an answer choosing path and the proportion of the number of correct choices in the length of the answer choosing paths. Among them, the length referred to the times of a student choosing an answer, and the proportion referred to the ratio of the total number of correct answers to the total length of the
answer choosing paths.

3.2.3. **Cluster analysis**

The following formula was used for cluster analysis in the paper.

\[
sk_{tx} = \frac{(s_{q_1} + s_{q_2} + \cdots + s_{q_v})}{v}
\]

Based on this formula, the student-knowledge point learning characteristic matrix SK was obtained.

\[
SK = \begin{bmatrix}
sk_{11} & sk_{12} & \cdots & sk_{1n} \\
sk_{21} & sk_{22} & \cdots & sk_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
sk_{h1} & sk_{h2} & \cdots & sk_{hn}
\end{bmatrix}
\]

Where, \(sk_{tx}\) was the student \(S_t\)’s understanding degree of a knowledge point \(K_x\); \(sk_{tx} \in \{0,1\}\); \(h\) was the number of students; \(n\) was the number of knowledge point; the value of \(sk_{tx}\) was proportional to the student \(s_t\)’s understanding degree of a knowledge point \(K_x\). The value of \(sk_{tx}\) was determined by the average value of the student’s understanding degrees \(s_{qa}, s_{qb}, \ldots, s_{qo}\) of the test questions \(qa, qb, \ldots, qo\) (there were \(v\) test questions in total).

3.2.4. **Construction of the knowledge graph**

Each set of test questions was integrated with a certain number of knowledge points, and students presented certain learning characteristics in the process of answering them. The students of the same learning characteristics were assembled, and then a DBSCAN clustering algorithm was used to reasonably adjust the radius and the minimum number of points included to achieve automatic clustering of the students. If the above two parameters could be used to accurately and swiftly identify the characteristics of the students, then it would be proved that the feasibility of the DBSCAN algorithm in achieving the automatic analysis of the students' answer data and the automatic construction of the knowledge graph model.

After clustering the students using the DBSCAN algorithm, we obtained \(ckZX = \{ck_{z1}, ck_{z2}, \ldots, ck_{zn}, \ldots, ck_{zn}\}\), in which \(ck_{xz}\) referred to the learning characteristics corresponding to a student category \(c_z\) and a knowledge point \(k_x\), respectively.

A knowledge graph consists of the relationship between entities. The entities are the "nodes" in the knowledge graph, and the relationship between the entities is the "sides" in it. In the research, the entities were the students participating in the evaluation, and the relationship referred to the association between the students. The latter was obtained by analyzing the data generated during the evaluation on the students, and mining and analyzing the learning characteristics of them.

The learning characteristics of the students were confirmed through the cluster analysis. Among all learning characteristics, the students’ best and worst understanding degrees of knowledge points are the optimal manifestation of the characteristics. The following three factors are generally considered to extract the learning characteristics of students: 1) The best knowledge points of the students of the same learning characteristics are identical; 2) The worst knowledge points of the students of the same learning characteristics are identical; 3) Both are identical. Generally, the mean value of the differences in the learning stability of each group will be used to verify the authoritativeness and reasonableness of the model, thereby ensuring the accuracy of the constructed knowledge graph of students.

4. **Conclusion**

Each student has his/her own characteristics in learning every knowledge point. In past teaching, teachers obtained these learning characteristics through artificial analysis. Characterized by low accuracy and efficiency, this method fails to provide targeted guidance for students' learning. The intelligent education model in this research makes full use of computer information technology, applies data mining and analysis to determine the characteristics of students in learning certain knowledge points, and classify the students of same learning characteristics through cluster analysis to
build a knowledge graph model. This model can systematically analyze the problems of students in learning, and give targeted suggestions, so that teachers can be more aware of the students’ understanding degrees of knowledge points, and propose personalized guiding programs, which is conducive to improving student performance and teaching quality.

With the continuous advancement of information science and technology, intelligent education becomes increasingly important, playing a critical role in enhancing teaching quality and students’ learning effectiveness. The pace of researching intelligent education has been started, but it is just a small step. There is still a long way to go in the future. For example, there is more room for the improvement of data algorithms, the classification of students can be further refined, and even the mighty calculation power of computers can be taken advantages. Thus, it is necessary to keep enriching the database of student learning characteristics to make a personalized learning improvement plan for each student, rather than relying on teachers, students and parents to fumble.

Reference
[1] Li Wei, Zhong shaochun. Research on Wisdom Teaching Model of Primary School Mathematics Supported by Information Technology[J]. 2016,(11):128-133.
[2] Tang weiwei, Fan yaqin. Study on the Teaching Strategies of Smart Mathematics Classroom in Primary Schools Based on Network Learning[J]. 2015,(7):49-54.
[3] Ma xiangchun, Zhong shaochun. Research on the Support Model and Implementation Mechanism of Personalized Adaptive Learning System from the Perspective of Big Data[J]. 2017,(4):97-102.
[4] Tang weiwei, Wang mengxue. Research on Intelligent Teacher Training Mode in Blended Learning Environment[J]. Education Research, 2015,(8):8-12.
[5] Pang jinwen, Gao linlin. Research on the Countermeasures of Informatization Leadership Training for Primary and Middle School Principals in the Mixed Learning Environment[J]. 2016,(6):20-27.
[6] Yang xianming, Yu shengquan. Smart Education System Framework and Key Supporting Technologies[J]. China Audio-visual Education, 2015,(1):77-84.
[7] Wang yuxi, XU da. Design of Inquiry Teaching Mode Based on Electronic Schoolbag—Take Primary School Science Teaching as an Example[J]. China Audio-visual Education, 2014,(2):95-100.
[8] Ma xiangchun, Zhong shaochun. Practical Research on the Flipped Classroom Teaching Mode Based on the Electronic Schoolbag Teaching System[J]. Research on Audio-visual Education, 2017,(6):111-115.
[9] Ma ling. Research on Classroom Teaching Innovation in Vocational Education from the Perspective of Wisdom Education[J]. Vocational Education Forum, 2017(011):90-92.
[10] Ma xiangchun, Zhong shaochun, Xu da. Research on the Support Model and Implementation Mechanism of Personalized Adaptive Learning System from the Perspective of Big Data[J]. China Audio-visual Education, 2017(4):97-102.