

Research on Data Mining and Value Implementation Path of University Student Archives

Shouzhong Wang^{1*}

¹ Archives, Henan University, Kaifeng, Henan, China

Email Address

wsz@henu.edu.cn (Shouzhong Wang)

*Correspondence: wsz@henu.edu.cn

Received: 29 August 2023; Accepted: 20 September 2023; Published: 23 October 2023

Abstract:

With the arrival of the information age, the accumulation and management of student record data in colleges and universities has become an important topic. The purpose of this paper is to study the implementation path of student record data mining in colleges and universities, and to explore its value for education management, personalised development of students and decision-making in colleges and universities. Firstly, the value of student record data mining in colleges and universities is analysed, including improving the efficiency of education management, assisting students' personalized development, and supporting decision-making and policy formulation. Then, the path research is carried out in three aspects: data collection and integration, data mining model selection and establishment, and data mining results interpretation and application. Finally, the results of the study are summarised and an outlook for future work is given.

Keywords:

University Student Records, Data Mining, Educational Management, Personalised Development, Decision Making

1. Introduction

With the rapid development of social economy and the continuous expansion of the scale of higher education, colleges and universities are faced with an increasing number of student record information processing tasks. The traditional manual processing has been unable to meet the needs of large-scale data management and application, so the introduction of data mining technology has become one of the solutions. As an emerging research field, data mining of student records in colleges and universities has a broad application prospect and far-reaching significance.

2. The Value Realized by Data Mining of Student Records in Higher Education

2.1. Enhancing the efficiency of student education and management in higher education

Data mining of student records in colleges and universities can effectively improve the efficiency of student education management in colleges and universities. Traditional manual processing cannot meet the needs of large-scale data management and application, and the introduction of data mining technology can help colleges and universities better understand and manage student information, thus improving the efficiency of education management. By mining and analysing student record data in colleges and universities, it is possible to more accurately understand students' personal information, learning situation and behavioural characteristics. Education administrators can use this information to develop personalised teaching plans and provide targeted guidance and counselling for each student. For example, by analysing a student's performance and course selection, learning difficulties or strengths in subjects can be identified and targeted resources and support can be provided. This will not only help students overcome difficulties and improve their performance, but also enhance their satisfaction and engagement.

Data mining technology can also help to discover the patterns and trends hidden in massive data. By integrating and analysing the information on students' performance and graduation destinations from different batches, majors, grades and other groups over the years, some important conclusions can be drawn, such as the common characteristics of outstanding students and the employment trends of different majors. These conclusions can provide an important reference basis for education administrators, helping them to better formulate teaching plans, adjust curriculum and improve education quality. In addition, data mining technology can also help colleges and universities discover some potential problems and take timely measures to intervene.

Data mining of student records in colleges and universities is of great significance in enhancing the efficiency of student education and management in colleges and universities. Through the use of data mining technology, it is possible to more accurately understand the situation of students, discover patterns and problems, and take appropriate measures to intervene in a timely manner. This will greatly improve the efficiency of education management in colleges and universities and promote the growth and development of students.

2.2. Supporting the personalised development of students in higher education

Data mining of student records in colleges and universities can assist in the personalised development of students in colleges and universities. Each student has different interests, abilities and career plans, and data mining technology can help colleges and universities better understand the characteristics and needs of students and provide them with personalised development support. Firstly, by mining and analysing data on students' personal information, interests and ability levels, it is possible to understand students' strengths and potential advantages. This information can be used to develop personalised course selection recommendations and elective directions. For example, by analysing students' course selection records and performance, it is possible to find out that they have a strong interest and potential in certain fields. Based on this information, colleges and universities can provide them with more suitable courses or directions of specialisation, as well as corresponding training and support, in order to promote their in-depth development in the field. Secondly, data mining technology can also help colleges and universities provide students with personalised career planning advice. By analysing data such as the employment situation of previous graduates and the employment trend in the industry,

it is possible to predict the impact of different major or course choices on future employment. In this way, colleges and universities can provide students with more targeted career counselling and employment guidance to help them find a more suitable career direction in the job market. In addition, data mining technology can also help colleges and universities identify difficulties and problems that students may encounter in the learning process and take timely measures to support them. By analysing data such as students' assignments, test scores, and records of course selection, it is possible to find out that students have greater difficulties in certain subjects or fields [1].

Data mining of student records in higher education is of great significance in assisting the personalised development of students in higher education. By using data mining technology to analyse students' personal information, interests and ability levels, etc., they can be provided with personalised course recommendations, career planning advice and learning support. This will help each student find the right direction for their development and achieve personalised growth and success.

2.3. Supporting decision-making and policy development in higher education

Data mining of student records in colleges and universities can support decision-making and policy formulation in colleges and universities. By mining and analysing the data of student records, it can provide important reference basis for the decision makers of colleges and universities, so as to make more scientific and reasonable decisions and policies. By mining and analysing the data on students' employment destinations and graduation, the impact of different majors or courses on employment can be understood. Such information can help universities better adjust their enrolment plans and major settings to meet the demands of the employment market. For example, if the employment rate of a major is low, universities can take corresponding measures to improve the educational quality of the major or adjust its enrolment scale. This will enhance the overall competitiveness of the university and better meet the needs of society. By mining and analysing data such as students' grades and course selection records, the impact of different courses or teaching methods on students' performance can be found. Such information can be used to improve the quality of teaching and teaching methods, as well as to help universities optimise their curriculum and allocation of teaching resources. For example, when it is found that the general performance of a certain course is low, universities can target to strengthen the teaching quality of the course, or improve the learning effect of students by adjusting teaching methods and resource allocation [2]. In addition, data mining technology can also help colleges and universities to identify potential problems and make predictions, through the analysis of student behavioural data, it can be found that there may be problems such as disciplinary actions, dropout risk and other problems, and timely take appropriate measures to intervene. This can effectively maintain the order and stability of colleges and universities and guarantee a good atmosphere in the educational environment.

Data mining of student records in colleges and universities is of great significance to support decision-making and policy formulation in colleges and universities. By using data mining technology to analyse information on students' employment, performance and behavioural characteristics, it can provide scientific and accurate references for university decision-makers. This will help to formulate more reasonable enrolment plans, employment policies and curricula, and enhance the overall competitiveness and quality of education in colleges and universities.

3. Research on the Path of Data Mining of Student Records in Colleges and Universities

3.1. Data collection and integration

Data collection and integration of student records in colleges and universities is an important part of education management in colleges and universities. It involves the collection and integration of student-related information to provide basic support for subsequent analysis and application. Data collection refers to the process of collecting student-related information. Such information includes basic personal information (e.g., name, gender, date of birth, etc.), academic information (e.g., year of enrolment, major, class, etc.), records of grades, course selection, and other data related to student development. Data collection can be carried out in a variety of ways, such as filling out online forms, system entry, and questionnaires. Data collection needs to be done in a way that ensures the accuracy and completeness of the data and adheres to relevant privacy protection policies. Data integration is the process of merging and unifying data from different sources and formats. Colleges and universities usually have multiple departments or systems responsible for different aspects of student management, so it is necessary to integrate these dispersed data into a unified database or platform. This can facilitate the subsequent comprehensive analysis and application of student information. Data integration needs to take into account the differences between data structures and formats, and carry out appropriate conversion and cleaning to ensure the quality of the integrated data.

When carrying out data collection and integration, it is also necessary to pay attention to the following points: (1) Data security: when handling students' personal information, it is necessary to comply with relevant data security and privacy protection regulations to ensure the security and confidentiality of students' information. Colleges and universities should establish a sound information management system and take necessary technologies and measures to prevent data leakage and misuse. (2) Data quality control: In the process of collecting and integrating data, data quality control is required, including the checking and cleaning of data accuracy, completeness and consistency. Universities can use automated tools or specialised personnel to carry out data quality checks to ensure data accuracy and reliability. (3) Data standardisation: in order to facilitate subsequent analysis and application, data from different sources need to be unified and standardised to ensure that data can be effectively compared and correlated with each other. Colleges and universities can establish unified data standards and convert and map the collected data accordingly. (4) Data update and maintenance: student information is dynamic and needs to be updated and maintained in time. Colleges and universities should establish a sound mechanism to ensure the timely update of data and solve problems such as data errors or conflicts. It is also necessary to regularly back up and archive student record data to prevent accidental loss or damage [3].

Data collection and integration of student records in colleges and universities is an important part of education management in colleges and universities. By effectively collecting and integrating student-related information and following the corresponding norms and processes, it can provide a reliable basis for subsequent analyses, decision-making and policy formulation. At the same time, colleges and universities also need to pay attention to data security and quality control to ensure the security, accuracy and completeness of student information.

3.2. Selection and building of data mining models

When performing data mining on college student records, it is crucial to choose a suitable data mining model. Different models have different characteristics and applicable scenarios, so it is necessary to select and build suitable models according to specific needs and data characteristics. Firstly, commonly used data mining models include classification models, clustering models, association rule models and prediction models. Classification models can classify students according to a certain standard, for example, classifying students into excellent, good, passing and other levels; clustering models can group students according to a certain similarity; association rule models can discover the correlation and patterns between students; prediction models can predict the future development trend of students through historical data. According to the specific needs, choose the appropriate data mining model for analysis. Secondly, the correlation between data volume, dimensions and features also need to be considered when choosing models. If the data volume is large and the dimension is high, neural network models based on deep learning algorithms can be considered; if the data volume is small and there is obvious correlation between the features, traditional machine learning algorithms such as decision trees or support vector machines can be used. In addition, the requirements for the interpretability of the results and the complexity of the model should be taken into account. Some models have better interpretability and can help understand student behaviour and patterns, while others may be more complex but provide more accurate prediction results. Choose the appropriate model based on specific needs and usage scenarios. Finally, steps such as data preprocessing, feature selection and model training are needed to build data mining models. Data preprocessing includes data cleaning, missing value processing, outlier processing, etc.; feature selection is to select the features that have an important impact on the target variables from a large number of features; model training is to estimate and optimise the parameters by inputting known data into the model. Building an effective data mining model requires several iterations and adjustments to obtain better performance [4].

It is very important to choose a suitable model in data mining of student records in colleges and universities. According to the specific needs and data characteristics, choose a suitable classification, clustering, association rules or prediction model, and build an effective data mining model through the steps of data preprocessing, feature selection and model training.

3.3. Interpretation and application of data mining results

The results of data mining of student records in colleges and universities can provide important references and support for educational management and decision-making in colleges and universities. These mining results need to be interpreted and applied appropriately to achieve better student development and education quality improvement. Firstly, interpreting the data mining results needs to be combined with specific analysis objectives and problem background. For the results of classification models, in-depth analysis can be conducted based on the characteristics of student groups and to understand the characteristics and differences of different categories of students; for the results of clustering models, different types of student groups can be identified and their commonalities and differences can be understood; association rule models can reveal the correlations and behavioural patterns of students; and prediction models can use historical data to predict future trends in student performance, graduation and other trends. By interpreting these results, student behaviour and

trends can be better understood. Second, the application of data mining results requires the development of appropriate strategies and measures for specific situations. For example, after identifying outstanding student groups in the classification model, colleges and universities can provide more high-quality resources or special cultivation programs for these students; if a certain category of students is found to be at higher risk in the clustering model, targeted counselling and intervention measures can be taken; the results of the association rule model can provide colleges and universities with guidance for improving student behaviour, for example, if certain behaviours are found to be related to academic performance, the cultivation of related behaviours can be strengthened. The cultivation of related behaviours can be strengthened. By combining data mining results with actual education management, more targeted and effective education strategies can be developed. Finally, data mining results can also be used to assess and monitor the quality of higher education. By analysing the indicators of student performance and graduation in the mining results, universities can understand the strengths and weaknesses of their own education quality, and make timely adjustments to teaching methods, curriculum and other aspects in order to improve the overall quality of education.

In conclusion, the interpretation and application of the data mining results of student records in colleges and universities need to be carried out in combination with specific problems and backgrounds, and corresponding strategies and measures should be formulated. Combining data mining results with actual education management can help colleges and universities better understand the characteristics of the student population, predict the development trend of students, and provide a scientific basis for improving the quality of education.

4. Conclusions

Data collection and integration of student records and data mining in colleges and universities is an important means to improve the level of education management and student development in colleges and universities. By effectively collecting, integrating and analysing student-related information, we can better understand the characteristics and behavioural patterns of the student population, and provide colleges and universities with scientific basis and decision-making support. When collecting and integrating data, it is necessary to focus on data quality control, security protection and standardised processing to ensure the accuracy, completeness and reliability of the data. It is also necessary to pay attention to the privacy protection of students' personal information and comply with relevant regulations and policies. Selecting appropriate data mining models and interpreting and applying their results can help universities gain a deeper understanding of student development trends, optimise the allocation of educational resources, and formulate appropriate educational strategies. This can better serve the development of students and improve the quality of education, and promote the development of university education in a more intelligent and personalised direction.

Conflicts of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

Funding

This work was supported by the Henan Provincial Archives Bureau “Research on the upgrading strategy of “digitalization” of archive information in colleges and universities” (NO.2020-X-07)

References

- [1] Jun, X. Research on the collaborative integration path of university archive departments under the governance environment of education big data. *Lantai World*, 2021, 8, 3. DOI:10.16565/j.cnki.1006-7744.2021.08.15.
- [2] An Buwing Miao Miao Yang Yuzhi. Research on the Value and Path of Red Archives in Colleges and Universities to Enhance the Identity of College Students as Party Members--Taking University S as an Example. *Shanxi Archives*, 2022, 3, 81-88.
- [3] Guo, M. Research on the realisation path of cultural value of university archives. *Popular Standardisation*, 2020, 21, 2.
- [4] Li, L.F. Research on Data Mining and Information Service of College Archives. *Office Business*, 2022, 22, 111-113.



© 2023 by the author(s); licensee International Technology and Science Publications (ITS), this work for open access publication is under the Creative Commons Attribution International License (CC BY 4.0). (<http://creativecommons.org/licenses/by/4.0/>)