Transactions on Social Science, Education and Humanities Research
ISSN: 2960-1770, eISSN: 2960-2262 | Volume 13
ECSS 2024

Data mining-based personalized learning manual customization for smart classrooms

Jingyang Liu

Hefei No.4 High School, Anhui 230000, China

Abstract. With the development of intelligent education, personalized learning has become an important way to improve students' learning efficiency and effectiveness. In this paper, a method for customizing personalized learning manuals for smart classrooms is proposed based on data mining technology. First, students' learning behavior data are collected through the smart classroom and preprocessed and feature extracted. Then, data mining technology is used to analyze and classify students' learning behaviors, so as to achieve accurate identification of personalized learning needs. Based on the analysis results, this paper designs a personalized learning manual generation system, which can dynamically adjust the learning content and path according to students' individual characteristics and learning progress. Finally, the effectiveness of the method is verified through practical application cases, and the results show that the customized personalized learning manual significantly improves students' learning experience and learning outcomes. The research in this paper provides new ideas and practical references for personalized teaching in smart classrooms.

Keywords: data mining, smart classroom, personalized learning, learning manual customization, learning behavior analysis, education technology.

1. Introduction

With the rapid development of information technology, smart education has gradually become an important part of modern education. With the help of a variety of advanced technological means, such as big data, artificial intelligence and the Internet of Things, smart classrooms provide educators with more accurate and efficient teaching tools[1]. In this context, personalized learning has emerged as an important way to meet the diverse learning needs of students and improve learning results. The traditional education model often fails to take into account the individual differences of each student, resulting in unsatisfactory learning results. By focusing on the unique needs and learning styles of each student, personalized learning can better stimulate students' interest and potential in learning[2].

Data mining technology, as a tool for extracting valuable information from massive data, has great potential for application in personalized learning[3]. The large amount of students' learning behavior data accumulated in the smart classroom provides rich materials for data mining, and through the analysis of these data, we can gain a deeper understanding of students' learning habits, knowledge mastery, and learning progress, so as to provide a scientific basis for the customization of personalized learning manuals[4].

The purpose of this paper is to explore how to customize personalized learning manuals in smart classrooms using data mining techniques[5]. By collecting and analyzing students' learning behavior data, this paper proposes a method for dynamically generating learning manuals, which can provide students with personalized learning resources and suggestions according to their actual needs and learning paths, thus improving learning efficiency and effectiveness[6]. The research in this paper not only provides theoretical support for personalized teaching in the smart classroom, but also provides reference for educators to apply personalized learning manuals in practice.

2. Overview of Smart Classrooms and Personalized Learning

Smart classroom is a product of the combination of modern educational technology and teaching practice, which utilizes advanced information technology, such as artificial intelligence, big data and the Internet of Things, to build a more flexible and interactive teaching environment[7]. In the smart classroom, teachers can accurately grasp each student's learning progress and weaknesses through real-time monitoring of students' learning, so as to make targeted adjustments to teaching strategies[8]. At the same time, the application of multimedia and virtual reality technology in the smart classroom makes the teaching content more vivid and the students' learning experience more rich. However, although the smart classroom provides strong support on the technical level, how to truly transform these technologies into a means to improve students' learning results is still a topic that requires in-depth research. Learning Behavior Feature Vector Extraction:

$$f_i = [f_{i1}, f_{i2}, \dots, f_{in}] \tag{1}$$

In the traditional mode of teaching, the content and pace of teaching are often uniform, making it difficult to take into account the individual differences of each student[9]. This "one-size-fits-all" approach fails to meet the diverse learning needs of students, resulting in some students losing interest in learning because they cannot keep up with the pace, while others fail to realize their full potential due to a lack of challenge. Personalized learning is the solution to this problem, through the analysis of students' individual characteristics and learning habits, tailored to their learning content and path, so as to improve the relevance and effectiveness of learning[10]. With the advancement of education technology, the need for personalized learning has become more and more urgent, especially in the scenarios of large-class teaching or online education. Prediction Model for Learning Outcomes:

$$\hat{y} = \beta_0 + \sum_{j=1}^{n} \beta_j f_j + \epsilon \tag{2}$$

As an important tool to support personalized learning, personalized learning manuals can provide customized learning resources and guidance according to students' learning needs. Through personalized learning manuals, students can obtain learning materials that match their learning levels and interests, make reasonable learning plans, and obtain timely feedback and suggestions. This kind of targeted learning guidance can not only help students overcome their learning difficulties, but also stimulate their motivation and independent learning ability. In addition, personalized learning manuals can also help teachers better understand students' learning progress so that they can make more effective teaching interventions.

The smart classroom provides technical support and data foundation for the implementation of personalized learning. Through smart devices and platforms, teachers can collect and analyze students' learning data in real time and develop personalized learning manuals accordingly. In the smart classroom, personalized learning is no longer a theoretical concept, but a concrete practice that can be implemented on the ground through data mining and other technical means. Through the combination of smart classroom and personalized learning, educators can not only improve the accuracy and efficiency of teaching, but also better meet the learning needs of each student, so as to achieve the true meaning of tailored teaching.

3. Data Mining Based Learning Behavior Analysis

Learning behavior analysis based on data mining mainly consists of three key steps: first, through the collection and preprocessing of learning data, we can obtain and organize valuable learning behavior data; second, through the feature extraction of learning behavior, we can identify and analyze various behavioral characteristics exhibited by students in the learning process; lastly, through the classification and prediction of these characteristics, we can accurately judge the Finally, through the classification and prediction of these features, we can accurately determine the learning status of students and provide a scientific basis for the customization of personalized learning manuals. These steps together constitute the complete process of learning behavior analysis, which provides important support for personalized teaching in the smart classroom.

3.1. Learning data collection and pre-processing

The large amount of data generated in the smart classroom provides rich material for learning behavior analysis. These data mainly come from a variety of learning activities and interactions in the classroom, including students' class participation, homework submission records, test scores, class discussion records, and online learning behavior data generated through learning management systems (LMS) or other educational platforms. These data types are rich in both structured data, such as grades and time records, and unstructured data, such as class notes and discussion texts. These diverse data provide the basis for a comprehensive analysis of student learning behavior, but they also increase the complexity of data processing.

In order to effectively collect learning data, smart classrooms typically employ a variety of technologies and tools. Sensor technologies, learning management systems (LMS), educational apps, and data collection platforms are common means of data collection. For example, cameras and microphones in the classroom can record students' classroom performance, LMSs can track students' progress and homework completion, and mobile learning apps can provide data on students' learning outside the classroom. These technologies not only automate data collection, but also update and synchronize data in real time, thus ensuring that learning data is timely and accurate. At the same time, these data collection tools are usually combined with big data platforms to form a complete data collection and storage system, laying the foundation for subsequent data mining.

Collected learning data often contain a lot of noise and incomplete information, so it is important to clean and preprocess the data before analyzing it. The main tasks of data cleaning include removing duplicate data, fixing or removing missing values, correcting errors in the data, and standardizing the data format. In a smart classroom scenario, noisy data may come from device malfunctions, non-learning behaviors of students (e.g., random clicks), or system error logs. With proper data cleaning and noise processing, the quality of data can be significantly improved to ensure the accuracy and reliability of subsequent analysis,showed in Figure 1:

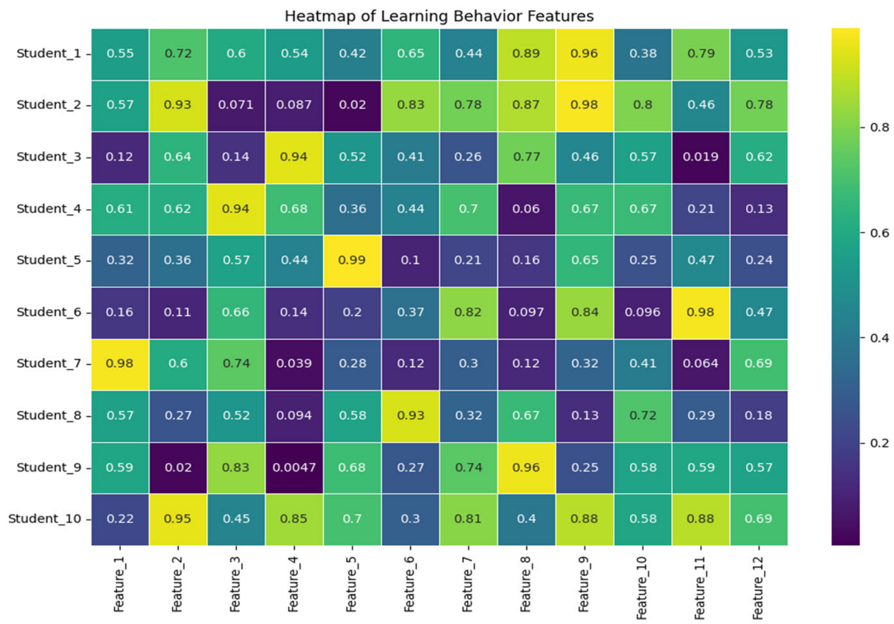

Figure 1. Heatmap of Learning Behavior Features

In the preprocessing stage, the selection and dimensionality reduction of data features is also a crucial step. Learning data generated by smart classrooms are usually high-dimensional and contain many irrelevant or redundant features, and analyzing them directly will not only increase the computational complexity, but also may lead to overfitting of the model. Therefore, it is necessary to filter out the key features that are closely related to learning behaviors, such as learning time, participation, correct answer rate, etc., through feature selection techniques. In addition, the dimensionality of the data can be reduced through dimensionality reduction techniques, such as Principal Component Analysis

(PCA), thereby reducing the computational cost while retaining the key information. This process helps to improve the efficiency and effectiveness of subsequent data mining algorithms, making the customization of personalized learning manuals more accurate.

3.2. Feature Extraction for Learning Behavior

Feature extraction of learning behaviors is a key step in the data mining process, aiming to extract core features from raw learning data that can represent students' learning status. These features can reflect various aspects of students' performance in the learning process, such as the frequency of learning, concentration, and knowledge mastery. By extracting and analyzing these features, educators can gain a deeper understanding of students' learning habits, strengths and weaknesses, thus providing a scientific basis for the customization of personalized learning manuals. The ultimate goal of feature extraction is to transform complex learning behavior data into indicators that are easy to analyze and interpret, thus supporting accurate behavior classification and prediction,showed in Figure 2 :

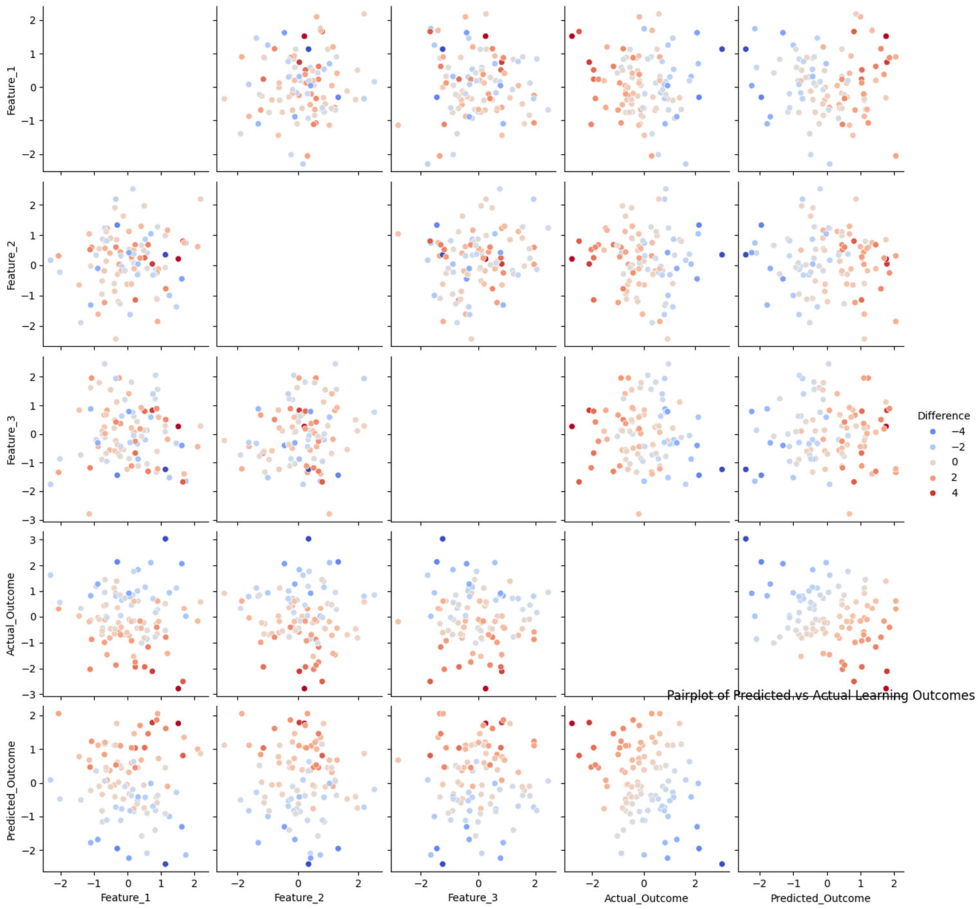

Figure 2. Pairplot of Predicted vs Actual Learning Outcomes

In the smart classroom, common learning behavior characteristics can be divided into several categories: time characteristics, interaction characteristics, performance characteristics and emotional characteristics. Temporal features include the distribution of learning time, learning duration, and learning frequency, etc. Interactive features cover the interaction between students and the teaching content, teachers, and other students, such as the number of questions asked, the number of times they participate in the discussion, etc. Performance features usually refer to the academic performance of students, including test scores, completion of homework, and mastery of knowledge points; affective features analyze the data of students' expressions, voice intonation, etc. to infer their learning emotions and psychological state. to infer their learning mood and psychological state. These characteristics can not only be analyzed individually, but can also be combined to form more complex behavioral patterns, providing rich information for the comprehensive analysis of learning behavior.

Feature extraction can be done using a variety of methods, the choice of which depends on the nature of the data and the goal of the analysis. For structured data, such as learning time and grades, basic features can be extracted using traditional statistical analysis methods; for unstructured data, such as classroom discussion texts and videos of students' expressions, natural language processing (NLP) and computer vision techniques are required. For example, with NLP, the topics and emotional tendencies of students' concerns can be extracted from classroom discussions; with computer vision, students' facial expressions and gestures in the classroom can be analyzed to infer their engagement and emotional states. In addition, feature engineering techniques are also commonly used to combine, transform and optimize features to enhance feature expressiveness and analysis.

After feature extraction, feature selection and noise reduction are usually required to ensure the simplicity and effectiveness of the model. Since the extracted features may be excessive or even contain redundant or noisy information, the most representative features need to be filtered out by feature selection techniques. Commonly used feature selection methods include filtering, wrapping, and embedding, which can retain the features that contribute most to the classification and prediction of learning behaviors by evaluating the relevance or importance of the features to the target variables. Meanwhile, for high-dimensional data, noise reduction and dimensionality reduction techniques such as Principal Component Analysis (PCA) can help to remove noisy information from the data and reduce the number of features, thus improving the efficiency of data mining algorithms and the generalization of models. This step ensures that the final feature set used for learning behavior analysis has high explanatory and predictive power.

3.3. Classification and Prediction of Learning Behavior

Learning behavior classification is a crucial step in the process of data mining. By classifying students' learning behaviors into different categories, educators can better understand students' learning status and adopt appropriate teaching strategies. Common learning behavior classifications include classifying students as "active participants", "passive learners" and "marginal learners". This classification not only helps to identify students who need special attention, but also provides a basis for the development of personalized learning manuals. In order to achieve accurate classification of learning behaviors, various machine learning algorithms, such as K-means clustering, decision trees, and support vector machines (SVMs), are often used. These algorithms can automatically categorize students into appropriate learning behavior categories by pattern recognition of their feature data.

In learning behavioral classification, model training and evaluation are critical steps to ensure the accuracy of classification results. First, the dataset needs to be divided into a training set and a test set so that its performance can be evaluated during the model training process. During the training process, suitable algorithms are selected and their parameters are adjusted so that the model can effectively recognize different classes of learned behaviors. Once the model training is complete, the model is evaluated using the test set, which usually employs metrics such as accuracy, recall, and F1 score to measure the model's performance. In addition, the stability and generalization ability of the model can be further verified by methods such as cross-validation. Through continuous iterative optimization, a model that can reliably classify learning behaviors is finally obtained to support subsequent personalized teaching.

Learning behavior prediction is a further application based on classification models, aiming to predict students' future learning performance or possible learning difficulties by analyzing their current learning behavior characteristics. The results of prediction can help teachers identify potential learning problems in advance and take preventive measures. For example, by predicting a student's likely test scores or homework completion, teachers can intervene before a student's performance slips, thus improving teaching and learning outcomes. Commonly used prediction methods include time series analysis, regression analysis, and deep learning models, which are able to handle different types of data and provide effective prediction results for different educational scenarios.

Prediction results not only need to be accurate, but also need to be easy to interpret and apply. For educators, complex mathematical models or black-box algorithms are difficult to use directly in actual teaching and learning, so in practical applications, it is often necessary to visualize and explain the prediction results. For example, trends in student learning behaviors and possible outcomes of predictions can be shown by generating easy-to-understand charts or reports. At the same time, the prediction results should also be integrated with teaching objectives and strategies to provide teachers with practical suggestions and guidance. By applying the results of classification and prediction to the customization of personalized learning manuals, each student can be provided with more precise learning resources and personalized guidance, thus maximizing learning outcomes.

4. Generation and application of personalized learning manuals

Personalized learning manuals are designed to meet the unique learning needs and goals of each student. Design principles should include relevance, flexibility and adaptability. Relevance means that the learning content, activities and resources in the handbook must match students' learning goals, interests and ability levels; flexibility is reflected in the ability of the handbook to adapt to student progress and change, allowing for adjustments to be made based on real-time feedback; and adaptability emphasizes that the handbook should be able to be dynamically updated based on different stages of student learning and feedback data. By following these design principles, it can be ensured that personalized learning manuals can truly serve each student and improve learning outcomes.

The process of generating personalized learning manuals usually includes several steps of data analysis, content matching, manual generation and feedback adjustment. First, students' learning needs and knowledge gaps are identified by analyzing their learning data and behavioral characteristics. Then, based on the analysis results, suitable learning resources and activities, such as practice questions, reading materials, video tutorials, etc., are selected and matched into students' learning manuals. Then, automated tools or systems are used to generate personalized learning manuals to ensure that the organization and presentation of content meets students' needs. Finally, the effectiveness of the manuals is evaluated by collecting feedback from students and adjusting and optimizing them accordingly. This process ensures that the learning manuals are personalized and relevant.

Personalized learning manuals can be used in a variety of instructional scenarios, including classroom instruction, out-of-class tutoring, and online learning. Online learning platforms can use personalized learning manuals to provide students with customized learning paths and resources to help them stay productive and goal-oriented in their independent learning. In classroom teaching, teachers can adjust the content and methods of instruction based on students' personalized learning manuals to provide targeted tutoring and support. In out-of-class tutoring, personalized learning manuals can serve as a guide for students to learn independently, helping them learn at their own pace and needs. These application scenarios demonstrate the wide applicability and practical value of personalized learning manuals.

Assessment and optimization are key steps in ensuring the effectiveness of personalized learning manuals. Assessment can be done by analyzing students' learning outcomes, feedback and learning progress. For example, students' performance after completing the tasks in the manual can be checked regularly to assess their achievement of the learning objectives. At the same time, feedback from students and teachers on the manual is collected to understand the problems and suggestions for improvement in its practical use. Based on these assessment results, adjustments and optimization of the manual content are carried out, such as updating learning resources, adjusting learning paths and improving feedback mechanisms. This process helps to continuously improve the quality and adaptability of personalized learning manuals and ensure their effectiveness and usefulness in actual teaching.

5. Conclusion

This paper discusses the generation and application of personalized learning manuals for intelligent classrooms based on data mining, providing a comprehensive analysis and implementation framework from the collection and preprocessing of learning data, feature extraction of learning behaviors, classification and prediction of learning behaviors to the generation and application of personalized learning manuals. Through the application of data mining technology, valuable information can be extracted from massive learning data to gain an in-depth understanding of students' learning behaviors, thus providing a scientific basis for personalized learning.

The collection and pre-processing of learning data lays the foundation for the subsequent analysis, and the acquisition and organization of data through a variety of technical means ensures the quality of data and the accuracy of analysis. Through feature extraction, we are able to identify the characteristics of students' learning behaviors, and then classify and predict them to accurately grasp students' learning status and needs. The results of these analyses provide the basis for generating personalized learning manuals, so that the learning content and resources can accurately match the individual needs of students.

The generation and application of personalized learning manuals not only enhances learning effectiveness, but also increases students' interest and autonomy in learning. In practical application, personalized learning manuals can be dynamically adjusted according to students' learning progress and feedback, ensuring their continuous adaptability and effectiveness. Through continuous evaluation and optimization of the personalized learning manual, its quality can be further improved to meet the diverse needs of different students. The research and application of data mining-based personalized learning manuals for smart classrooms demonstrate the great potential of the field of smart education. In the future, with the continuous development of technology and the advancement of data mining methods, personalized learning manuals are expected to be widely used in more educational scenarios and make greater contributions to the personalized and intelligent development of education.

References

[1] Osman L, Chung A K K .UK national survey on personalized customization of A-constant in cataract surgery[J]. Eye (London, England), 2010, 24(5):938-940.DOI:10.1038/eye.2009.238.

[2] Growing with smart products: Why customization capabilities matter for manufacturing firms[J]. Journal of Product Innovation Management, 2023, 40(6):794-816.DOI:10.1111/jpim.12680.

[3] Lee J Y, Park J H, Kim H, et al. The FaaS system using additive manufacturing for personalized production[J].Rapid Prototyping Journal, 2018, 24(9):1486-1499.DOI:10.1108/RPJ-11-2016-0195.

[4] Schaefer D, Panchal J H, Choi S K, et al. Strategic Design of Engineering Education for the Flat World[J]. International Journal of Engineering Education, 2008, 24(2):274-282.

[5] Shi H, Yu J, Duan T. Advances in personalized modelling and virtual display of ethnic clothing for intelligent customization[J]. AUTEX Research Journal, 2024, 24(1):28-41.DOI:10.1515/aut-2023-0040.

[6] Wang Z, Tao X, Zeng X, et al. Design of Customized Garments Towards Sustainable Fashion Using 3D Digital Simulation and Machine Learning-Supported Human–Product Interactions[J]. International Journal of Computational Intelligence Systems, 2023, 16(1).DOI:10.1007/s44196-023-00189-7.

[7] Wang C, Zhu Y, Liu K Y J. Multifaceted Relation-aware Meta-learning with Dual Customization for User Cold-start Recommendation[J].ACM transactions on knowledge discovery from data, 2023, 17(9):1.1-1.27.

[8] Taleb T, Mada B, Corici M I, et al. PERMIT: Network Slicing for Personalized 5G Mobile Telecommunications[J]. IEEE Communications Magazine, 2017, 55(5):88-93.DOI:10.1109/MCOM.2017.1600947.

[9] Avancini H, Candela L, Straccia U. Recommenders in a personalized, collaborative digital library environment[J]. Journal of Intelligent Information Systems, 2007, 28(3):253-283.DOI:10.1007/s10844-006-0010-3.

[10] Makris C, Panagis Y, Sakkopoulos E, et al. Category ranking for personalized search[J].Data & Knowledge Engineering, 2007, 60(1):109-125.DOI:10.1016/j.datak.2005.11.006.