Research on the Data Analysis Scheme of Flexible And Personalized Curriculum Design of University Mathematics

Jian Zhang*, Cuiping Zhang

Department of General Education, Liaoning University of International Business and Economics, Dalian 116000, Liaoning, China

* Corresponding author: zhangjiant@luibe.edu.cn

Abstract: Flexible and personalized curriculum design can adjust teaching content in real time according to the learning situation feedback from students, which is crucial for promoting students' personalized curriculum development. Such design is highly dependent on the accurate grasp and dynamic response to students' learning situation. However, there are relatively few studies on flexible and personalized curriculum design schemes. Therefore, this paper constructs a set of schemes for learning situation collection, analysis, and flexible personalized curriculum design in university mathematics. It collects learning situation data from multiple scenarios, and after data cleaning, deduplication, and standardization, uses methods such as statistical analysis, knowledge graph, and machine learning to conduct learning situation analysis. Based on this, a flexible and personalized curriculum design scheme with hierarchical real-time adjustment is designed. In addition, this paper also provides rules and measures for the protection of students' private data. This study offers a feasible scheme for realizing precise and personalized teaching.

***Keywords*:** learning situation data; data collection and analysis; flexible and personalized course design

1. Introduction

University mathematics education serves as the core foundation for cultivating students' logical thinking, abstract modeling abilities, and scientific research literacy, providing rigorous theoretical support and analytical tools for innovative applications in various disciplinary fields. [1, 2]. Traditional university mathematics teaching is often conducted in large-class mode, making it difficult for teachers to take individual differences into account. Moreover, such a tendency towards standardized, mass-produced, and mechanized teaching restricts the cultivation of students' mathematical thinking and application abilities. The breakthroughs in Artificial Intelligence (AI) technology have offered a possibility to solve this predicament.

As an auxiliary tool, AI can replace teachers in completing repetitive teaching tasks such as explaining basic knowledge points and correcting homework, thus freeing up teachers' energy to focus on personalized curriculum design [3]. Precise education is realized by intelligently analyzing students' learning data, pushing customized resources, and providing real-time interactive feedback, enabling teaching students in accordance with their aptitude to move from an idea to practice [4, 5]. The core of achieving this goal lies in constructing a mechanism that can dynamically capture, analyze, and respond to students' learning status, which relies on efficient capabilities in data collection, filtering, analysis, and flexible personalized curriculum design [6].

To this end, this paper designs a flexible and personalized curriculum design scheme for university mathematics that adjusts the curriculum content in real time according to changes in students' learning situations. The scheme covers the stage of generating learning situation data, the stage of processing data (including collection, filtering, and analysis), and the stage of applying data in flexible and personalized curriculum design, and finally optimizes students' learning effects through feedback data. The schematic diagram of student learning data flow is shown in Figure 1. In addition, this paper also discusses methods for protecting students' privacy.

2. Data collection and filtering methods

2.1 Data collection methods

Data collection is the foundation of learning situation data analysis. Only by ensuring that the collected data covers all scenarios of learning, including pre-class, in-class, and after-class, can the comprehensiveness of learning situation analysis be guaranteed.

Pre-class preview data generally includes: data such as micro-video viewing trajectories, pause times, and replay

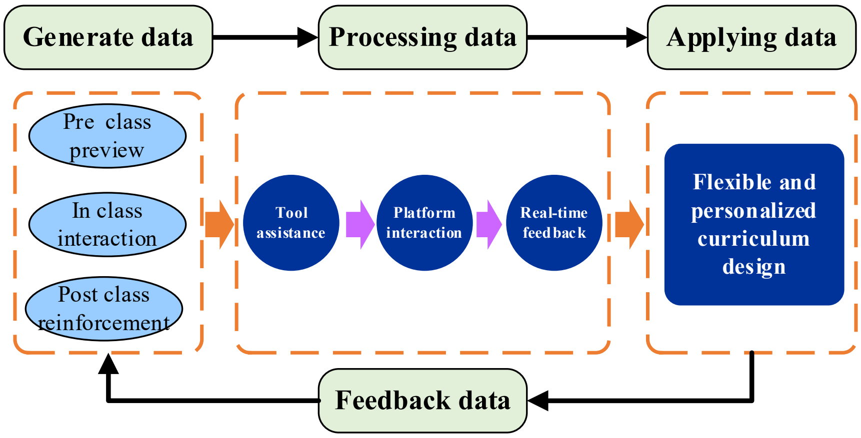

Figure 1. Schematic diagram of student learning data flow

segments collected through intelligent learning companion platforms; as well as online test data like answer time consumption, wrong options, and high-frequency words in modification records. In-class interaction data usually includes: information such as the accuracy rate of in-class tests during teaching, the number of mentions of key words in group discussions, and electronic whiteboard operation trajectories collected by means of intelligent classroom terminals. After-class consolidation data typically includes: problem-solving steps submitted through homework systems, marking notes, and question records on intelligent Q&A platforms. In addition, information such as students' majors, university entrance examination mathematics scores, and historical grades collected through the school system can be used to supplement learning background information.

2.2 Data filtering methods

The collected raw data generally cannot be directly used as final data for analysis; it needs to undergo multiple processing steps such as data cleaning, deduplication, and standardization to improve data validity. Among these, data cleaning can eliminate invalid records such as repeatedly submitted data and accidental clicks by students. Deduplication and merging integrate duplicate information. Standardization processes time data into a uniform format; for example, multiple-choice, fill-in-the-blank, and solution-based questions are normalized to values within the 0-1 interval, thereby facilitating cross-dimensional data analysis. The filtered data retains core information directly related to learning status. Compared with the raw data, the data volume can be reduced by 30%-50%, which can significantly improve the efficiency of subsequent analysis.

3. Precise data analysis methods

3.1 Data analysis methods

Common methods for analyzing learning situation data include statistical analysis, knowledge graph analysis, and machine learning algorithms. Among them, statistical analysis grasps the overall level of the class through descriptive statistics such as average score, scoring rate, and standard deviation. The knowledge graph analysis method constructs a three-layer graph of "knowledge points - question types - abilities". It locates students' knowledge gaps and provides replacement conditions for unmastered content. The machine learning algorithm uses a decision tree model to analyze the key factors affecting students' grades. In addition, K-means clustering is used to divide students into groups such as "computation-weak type", "concept-vague type", and "comprehensive excellent type", providing a basis for group teaching.

3.2 Data analysis software

Common data analysis software includes SPSS, Python, R language programs, etc. Among them, SPSS is suitable for basic statistical analysis with simple operations. It can quickly generate "bar charts of scoring rates for each chapter" and "knowledge point correlation matrices", which is convenient for teachers to intuitively grasp the overall situation of the class. However, it is difficult to handle unstructured data such as handwritten problem-solving steps. Python-based data analysis packages such as Pandas and Scikit-learn support the integration of multi-source data, enabling large-scale data cleaning and grade prediction. R language is more adept at statistical modeling and visualization, generating "time series curves of learning behaviors", but it has a relatively high threshold for teachers who are not majoring in statistics.

Overall, appropriate data analysis methods and software can be selected according to the characteristics of different stages of the course and different types of data to process complex learning situation data, thereby more accurately and deeply mining potential information in students' learning situation data.

4. Flexible and personalized curriculum design

Flexible and personalized curriculum design refers to the real-time adjustment of teaching plans according to the personalized learning situation feedback from students. This method realizes targeted curriculum design through a weighted fusion strategy, which comprehensively considers students' short-cycle learning situation, long-cycle learning situation and individual differences.

For students with weak foundations, the focus is on consolidating core concepts. Visual explanation resources are pushed, along with step-by-step exercises ranging from direct application of formulas to calculations combined with physical scenarios. A small test is arranged once a week, accompanied by targeted explanations. For problems with a high error rate, relevant practice content is added to gradually improve basic abilities.For students at the intermediate level, on the premise of mastering the basics, more comprehensive application questions are added. Cases with multiple solutions to one problem are pushed. Students are encouraged to participate in introductory mathematical modeling projects to cultivate their awareness of application.For top-performing students, extended content, interdisciplinary cases are pushed, and open-ended tasks are assigned. Study sessions led by top students are organized, where they explain their problem-solving ideas to strengthen high-order thinking.

Flexible and personalized curriculum design realizes multi-cycle learning situation comparison through the analysis of learning situations in long and short learning cycles, achieves real-time dynamic revision of curriculum plans, avoids the solidification of curriculum design caused by information cocoons, and ensures smooth channels for students' level improvement.

5. Rules for protecting students' privacy data

In the era of big data, data privacy is an unavoidable topic, and the protection of students' private data is inevitably involved in the process of learning situation data analysis. This study protects students' private information from four links: information collection, storage, use, and management.

In terms of the collection link, the "minimum necessity" principle is clearly defined, and irrelevant information such as ID numbers and home addresses is not collected; students are informed through pop-ups that "the data is only used for personalized teaching and will be strictly kept confidential", and the collection is carried out only after obtaining authorization. In the storage link, the AES-256 encryption algorithm is used to store data, and firewalls and intrusion detection systems are deployed on the server; data is backed up regularly, and backup files are physically isolated from the main files.In the use link, the data is anonymized, and the real names are replaced with "student ID + random code"; teachers can only access the aggregated data of the classes they teach and the individual analysis results of the students they supervise, and all operations are tracked with records of "who viewed which data at what time".

In terms of management systems, the Student Data Protection Specifications are formulated, clearly stating that "data is prohibited from being used for purposes other than teaching"; regular teacher training is carried out to strengthen the awareness of privacy protection; complaint channels are set up to promptly handle potential risks of data leakage. Table 1 shows the privacy protection measures for students.

Table 1. Privacy protection measures for students.

Link	Specific measures
Collection	Do not collect unrelated information; pop-up notification; authorized to collect
Storage	Use encrypted storage; deploy firewalls and intrusion detection systems; back up data regularly
Utilize	Data anonymization processing; Restricting teacher access permissions and operation records
Management	Formulate the “Student data protection standard”; strengthen the awareness of privacy protection

6. Conclusions

This study realizes the flexible and personalized curriculum design for university mathematics through three major modules: full-scenario data collection and filtering, precise analysis, and flexible personalized curriculum customization, and constructs protection rules for students' private data. In the future, it is necessary to optimize and develop the teacher-AI collaborative decision-making function to further promote the transformation of university mathematics education from standardization to personalization.

Acknowledgments

This is the phased research achievement of the 2024 Basic Scientific Research Project of Liaoning Provincial Department of Education, titled Research on the Value-Added Evaluation of College Students' Comprehensive Quality Based on Analytic Hierarchy Process (Project No.: LJ132410841013).

References

[1] Du, L., Xu, S., Xu, Z. Teacher-student-ai collaborative classroom: the role and practice of AI in empowering university mathematics education, Chinese Journal of Distance Education, 2025, (04): 59-65+81.
[2] Hu, Q.T., Wei, M., Ling, X.L., et al. Research on teacher classroom teaching behavior in the digitalization process: Historical origins, current dilemmas, and path choices. Chinese Journal of Distance Education, 2025,45(07), 83-97.
[3] Hao, N., Gui, Z. The technological models, real symptoms, and practical approaches of AI empowering "precise ideological education" in universities, Heilongjiang Higher Education Research, 2025, 43(07): 1-10.
[4] Jiang, H., Wang, X. C., Yang S.D. The potential applications, risks, challenges, and response strategies of generative ai in education, Modern Education Management, 2023, (07): 66-74.
[5] Wu, G. J. Post-human condition and chinese educational practices: education's end or lifelong education? — Philosophical reflections on education in the age of artificial intelligence, Journal of East China Normal University(Educational Sciences), 2019, 37(01): 1-15+164.
[6] Yuan, J., Wu, F. The shift in the logic of knowledge production in the age of AI and educational responses, Chinese Journal of Distance Education, 2025, 45(07): 20-34.