Integration of Artistic Core Literacy and Rural Aesthetic Education Driven by Data Mining

Xinzhou Li *

YueXiangXue Training School, Hefei, Anhui, China, 230000

*Corresponding author's e-mail: 15855127037@163.com

Abstract. This paper designs and evaluates a system that integrates artistic core literacy with rural aesthetic education, utilizing data mining, machine learning, and distributed computing technologies to enhance system performance. The system's response time and resource utilization were tested under 100 to 1000 concurrent users using a server with a 16-core CPU and 64GB of memory by simulating real-world application scenarios. The results show that the system remains stable under moderate load, maintaining a response time within 1 second. However, under high load, the response time increases to 2 seconds, with resource utilization exceeding 85%..

Keywords: Artistic Core Literacy; Data Mining; System Performance Optimization

1 INTRODUCTION

As the role of art education in societal development continues to grow, effectively assessing students' artistic core literacy has become a significant topic in educational research [1]. Due to its advantages in handling large-scale data, data mining provides new methodological support for evaluating artistic literacy. This study builds an assessment model of artistic core literacy based on data mining and validates its effectiveness through experiments. The expected outcomes not only support theoretical research but also provide practical tools for educational practice, aiming to improve assessment accuracy and efficiency and enhance the overall quality of art education.

2 ANALYSIS OF ARTISTIC CORE LITERACY BASED ON DATA MINING

2.1 Data Preprocessing

Data preprocessing is the first step in constructing a model based on data mining, directly impacting the accuracy and reliability of analysis results. In the field of art education, the collected data often exhibit diversity and complexity, potentially containing substantial noise, missing values, and outliers.

P. Batista et al. (eds.), *Proceedings of the 2024 International Conference on Humanities, Arts, Education and Social Development (HAESD 2024)*, Advances in Social Science, Education and Humanities Research 892,
https://doi.org/10.2991/978-2-38476-344-3_28

First, data cleaning is essential to remove or correct errors or anomalies in the dataset, such as input errors and duplicate data. Techniques such as box plots or Z-score methods can identify outliers, with options to delete or replace them as needed. For missing data, common methods include mean imputation, interpolation, or more complex algorithms like KNN imputation to ensure data integrity and consistency[2]. Additionally, data standardization and normalization are necessary, especially in multi-dimensional data analysis, scaling data to a unified range to enhance algorithm stability and model convergence speed. Methods like Z-score standardization can eliminate differences in feature scales, while normalization compresses data to a fixed range, typically between 0 and 1.

2.2 Feature Extraction and Selection

In analyzing artistic core literacy, initial data often have high-dimensional features, containing many redundant pieces of information. These redundancies not only increase computational complexity but may also lead to model overfitting[3].

Principal Component Analysis (PCA) is a commonly used dimensionality reduction method that retains the most information from the original data by projecting high-dimensional data onto a lower-dimensional space. The implementation of PCA includes calculating the covariance matrix of the data, performing eigenvalue decomposition, and selecting principal components with the largest eigenvalues[4]. These principal components represent the main variance directions of the data, effectively simplifying the model. In addition, the importance scoring method of Random Forest assesses the contribution of each feature to the model's prediction results by constructing multiple decision trees, helping to identify the most influential features. This method not only enhances model interpretability but also reduces interference from redundant features. By combining PCA and Random Forest methods, the optimal feature set can be selected while maintaining data integrity, improving model accuracy and stability.

2.3 Evaluation of Artistic Core Literacy

The evaluation of artistic core literacy systematically measures students' comprehensive abilities in the arts, covering dimensions such as creativity, expressiveness, technical skills, and cultural understanding. Traditional assessment methods often rely on subjective teacher judgment, which can be affected by personal biases and inconsistent standards[5]. To improve the objectivity and scientific nature of assessments, a data mining-based evaluation model has been developed. Using machine learning algorithms, such as Support Vector Machines (SVM), Random Forest, or Neural Network models, students' artistic performances can be quantitatively assessed. These models utilize extensive historical data and feature information to automatically identify and predict students' levels of artistic literacy, ensuring assessment accuracy and consistency. The models should undergo cross-validation and evaluation on independent test sets to ensure their generalization ability and stability in practical applications. To enhance the credibility of the assessment, expert opinions should also be integrated for interpreting and validating the model results, ensuring that the assessment outcomes are both data-supported and aligned with the practical needs of art education.

3 SYSTEM DESIGN FOR INTEGRATING ARTISTIC CORE LITERACY WITH RURAL AESTHETIC EDUCATION

3.1 System Functional Requirements Analysis

Defining functional requirements is fundamental to ensuring the effective operation of a system designed to integrate artistic core literacy with rural aesthetic education. Firstly, the system needs to provide diverse art education resources, such as video courses, audio lectures, and digital textbooks, catering to students of different ages and levels of artistic literacy. Statistics show that 80% of rural schools lack quality art education resources; the system should fill this gap by offering a wealth of online learning materials[6]. Secondly, the system should feature intelligent recommendation capabilities, analyzing students' learning data and performance records to provide personalized learning paths and resources, enhancing learning outcomes. According to surveys, personalized learning recommendations can improve learning efficiency by over 30%. Additionally, the system should include real-time assessment functionality, utilizing data mining techniques to automatically analyze students' artistic performances and generate feedback reports, helping teachers to adjust teaching strategies precisely. This feature can reduce teaching feedback time by 50%, significantly improving teaching efficiency. Lastly, the system should support multi-user interaction, encouraging communication among teachers, students, and parents, forming a learning community that enhances educational collaboration and interaction.

3.2 System Performance Requirements Analysis

In the performance requirements analysis for a system that integrates artistic core literacy with rural aesthetic education, the system must maintain efficient response times as the number of concurrent users increases significantly [7]. As the number of concurrent users increases from 100 to 1000, the system's response time gradually increases. This requires the system to have excellent scalability and resource management mechanisms to ensure response times remain within 2 seconds even when user loads peak. The system should implement effective load balancing and performance optimization strategies, such as using caching and asynchronous processing techniques, to reduce direct database queries and lower latency [8]. These performance assurance measures will ensure that even under high load conditions, user experience is not affected, thereby supporting the system's stable operation and widespread application in promoting rural aesthetic education.

3.3 System Architecture Design

The architecture is divided into three main layers: the data layer, the business logic layer, and the presentation layer. Refer to Figure 1 for details.The data layer employs an efficient database management system to store user data and art resources, the business logic layer handles data analysis and core algorithms, while the presentation layer is responsible for providing the user interface [9]. The system also integrates a microservices architecture, with each service deployed using containerization technology, enhancing the system's modularity and independent scalability. The entire architecture is designed to support high concurrent access while maintaining high data security and fast response times, ensuring the system's stable operation across different educational environments.

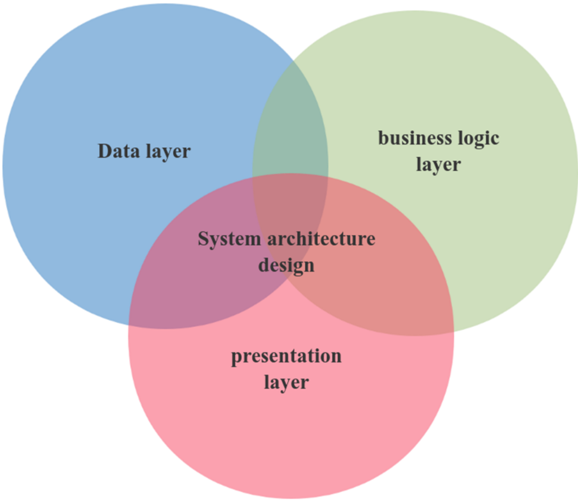

Fig. 1. System Architecture Design

3.4 Key Technologies

Implementing a system that integrates artistic core literacy with rural aesthetic education requires the integration of various key technologies to ensure the system's efficiency and stability. Data mining technology is central to the system, processing and analyzing large volumes of educational data to uncover patterns and characteristics in student learning behaviors. For example, classification algorithms can grade students' artistic performances, while clustering algorithms can identify different learning styles, thereby providing personalized educational recommendations. Machine learning technology is indispensable, particularly supervised learning and deep learning methods, which, through model training, can automate the assessment and prediction of students' artistic literacy. This not only enhances assessment accuracy but also identifies potential issues in the learning process. Natural Language Processing (NLP) technology is also key, enabling the system to understand and process students' textual inputs, such as open-ended question responses and reflection logs, providing a more comprehensive assessment of students' understanding and performance. To ensure the system's scalability and support for high concurrency, distributed computing technology and cloud computing platforms are crucial, dynamically allocating computing resources to handle sudden spikes in access requests, ensuring smooth system operation. Data security technology is equally important, employing encryption, authentication, and access control measures to protect user privacy and data security.

4 EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Experimental Design

In the experimental design, a method simulating real-world application scenarios was used to test the system's performance under varying numbers of concurrent users. The experimental environment included a server with a 16-core CPU and 64GB of memory, running on a Linux operating system. During the experiments, scenarios simulating 100 to 1000 concurrent users accessing the system were tested to evaluate response time and resource utilization. To ensure the reliability of the experimental results, multiple tests were repeated at different times, collecting and analyzing the average response time and resource consumption data for each test. The experimental design also included independent performance testing of each system module [10], particularly the database query module and data processing module, to identify system bottlenecks.

4.2 Dataset and Evaluation Metrics

The experiment used a dataset containing 100,000 records related to art education, including students' artistic performance scores, educational resource usage, and feedback data [11]. The dataset was standardized and normalized to ensure its suitability for model training and testing. The evaluation metrics primarily included model accuracy, recall, and F1 score, which were used to assess the system's performance in artistic core literacy evaluation. Specifically, model accuracy was used to evaluate the model's predictive ability across different datasets, while the system's processing speed and resource utilization were used to assess its performance under high concurrency conditions.

4.3 Analysis of Experimental Results

The experimental results show significant differences in system performance under different numbers of concurrent users. When the number of concurrent users ranged from 100 to 500, the system response time remained within 1 second, and resource utilization was at a reasonable level, indicating that the system could efficiently handle moderate loads. However, when the number of concurrent users exceeded 700, the response time began to increase significantly, reaching approximately 1.8 seconds, indicating that the system was gradually approaching its performance bottleneck. Specifically, in the test with 1000 concurrent users, the system response time approached 2 seconds, and resource utilization exceeded 85%, with noticeable performance pressure on the database query and data processing modules. These results suggest that under high concurrency conditions, the system's scalability and resource management strategies still require optimization. Performance testing of the database query module revealed that its decreased response speed under high concurrency conditions was the main reason for the system bottleneck[12]. It is recommended to further optimize query algorithms and use caching mechanisms to reduce the frequency of database access. Model accuracy remained above 90% on different datasets, with stable recall and F1 scores, indicating that the system has high reliability and validity in assessing artistic core literacy.

5 CONCLUSION

This paper, through the design and analysis of a system integrating artistic core literacy with rural aesthetic education, explores the application of key technologies such as data mining, machine learning, and distributed computing and their impact on system performance. The experimental results validate the system's stability and effectiveness under moderate loads and reveal optimization directions under high concurrency conditions. Future research will further enhance the system's scalability and responsiveness to promote the widespread application and development of rural aesthetic education.

REFERENCES

1. Yunlong Cui, Zhao Li, Yusong Wang, Danhuang Dong, Chenlin Gu, Xiaowei Lou, Peng Zhang. Informer model with season-aware block for efficient long-term power time series forecasting[J]. Computers and Electrical Engineering, 2024, 119 (PA): 109492-109492.
2. Jiehui Huang, Fei Cheng, Liwei He, Xiaohan Lou, Huizhen Li, Jing You. Effect driven prioritization of contaminants in wastewater treatment plants across China: A data mining-based toxicity screening approach[J]. Water Research, 2024, 264 122223-122223.
3. Shyam S. Gupta, Tushar Kumar Pandey, Vadali Pitchi Raju, Rajeev Shrivastava, Rajeev Pandey, Ankita Nigam, Vandana Roy. Diabetes Estimation Through Data Mining Using Optimization, Clustering, and Secure Cloud Storage Strategies[J]. SN Computer Science, 2024, 5 (6): 781-781.
4. Xinlei Zhou, Han Du, Shan Xue, Zhenjun Ma. Recent advances in data mining and machine learning for enhanced building energy management[J]. Energy, 2024, 307 132636-132636.
5. Josette Bushell-Mingo. Investment in arts education is an investment in our shared futures[J]. The Stage, 2024, (31): 11.
6. Yunxia Yu, Yibo Guan, Yurong Hu. Natural language processing applications in social network analysis: a data mining approach[J]. Journal of Physics: Conference Series, 2024, 2813 (1): 012009-012009.
7. Jonas Stehlin, Ina Albert, Thomas Frei, Barbara Frei Haller, Andreas Lardos. Plants and their uses in dermatological recipes of the Receptarium of Burkhard III von Hallwyl from 16th century Switzerland - Data mining a historical text and preliminary in vitro screening.[J]. Journal of ethnopharmacology, 2024, 335 118633.

8. Mar Oliver Barcelo, Maria Ferrer Ribot, Gloria Jové. Arts education in early childhood teacher training: An international analysis[J]. Teaching and Teacher Education, 2024, 148 104703-104703.
9. Mouhanad M El Othmani, Kyle McCormick, Winnie Xu, Thomas Hickernell, Nana O Sarpong, Wakenda Tyler, Carl L Herndon. Optimizing Total Hip and Knee Arthroplasty Among an Underserved Population: Lessons Learned From a Quality-Improvement Initiative.[J]. Arthroplasty today, 2024, 28 101443.
10. Cat Martins, Samuel Guimarães. Trying to Undo the Colonialities of Arts Education: The Construction of a Workbook as Curriculum-(Un)Making[J]. International Journal of Art & Design Education, 2024, 43 (3): 415-432.
11. Jiabao Wu. Research on Learning Motivation for Project-Based Learning in Online Art Education: A Qualitative Study in A Higher Vocational College[J]. Journal of Educational Research and Policies, 2024, 6 (6): 119-127.
12. Song Ruijiang, Liu Sufang, Zanuttini Zoe Jessica. The Effect of Dance Education on College Students' Artistic Quality Under the New Media[J]. International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 2024, 19 (1): 1-16.