# Research and Practice of Student Comprehensive Evaluation System Based on AI under the Background of Big Data

**Nie Wei[1], Yongjun Wei[2], Huanle Huang[1,*]**

[1]Liuzhou Vocational and Technical College, Liuzhou 545006, Guangxi
[2]Liuzhou Railway Vocational Technical College, Liuzhou 545616, Guangxi
**Correspondence Author*

**Abstract:** ***With the rapid development of big data and artificial intelligence technologies, the education field is undergoing unprecedented changes. The traditional student evaluation system generally relies on subjective evaluation and standardized tests, with drawbacks such as strong subjectivity and single evaluation dimensions. This research uses big data and artificial intelligence technologies to collect and mine college students' learning and living behavior data, constructs an index system for college students' comprehensive quality evaluation, and develops a student comprehensive evaluation system to provide decision-making basis for students' growth, education and teaching, and job selection and employment.***

**Keywords:** Big Data; AI; Student Comprehensive Evaluation; Evaluation Index.

## 1. INTRODUCTION

In the field of education, the evaluation of students' comprehensive qualities has always been an important topic in educational reform. The traditional evaluation system often focuses on students' academic achievements, neglecting the overall cultivation and development of students' comprehensive qualities. With the diversification of social talent needs, there is an urgent need for an evaluation system that can comprehensively reflect students' comprehensive qualities. The rise of big data and artificial intelligence technologies provides new ideas and methods for solving this problem. To explore a more scientific and objective way to evaluate college students' comprehensive qualities, relying on the construction of a smart campus, colleges and universities can gather a vast amount of data resources. This study integrates big data processing, knowledge graph construction, and artificial intelligence technologies, fully excavates the data value, gives play to the advantages of big data in data collection, storage, and efficient processing, uses deep learning technologies to deeply explore the characteristic patterns of students' behaviors, and meanwhile, with the help of the powerful analysis and reasoning ability of the knowledge graph, conducts intelligent analysis and exploration of students' behaviors based on a rich knowledge base. This process realizes the accurate identification and prediction of students' behavior patterns, providing support for risk early warning and management strategy formulation. Finally, these intelligent means provide strong support for student management personnel, significantly improving work efficiency and service quality, deepening the comprehensive understanding of students, making management work more goal-oriented and process-standardized, and thus better promoting the all-round development and healthy growth of students.

## 2. CURRENT SITUATION ANALYSIS

With the continuous progress of information technology, a large amount of data has been accumulated in the field of education, including students' academic achievements, classroom performance, participation in extracurricular activities, etc. These data provide rich resources for constructing a comprehensive student evaluation system. Foreign scholars' research on the academic performance of college students mainly focuses on the theoretical framework of students' academic performance, the data sources of academic prediction, effective academic indicators and models, the evaluation of the effect of prediction models, etc. From the perspective of the research on the comprehensive evaluation of college students, foreign research is relatively lacking, especially the research under the background of the big data era. Previously, the US Department of Education released a report titled "Promoting Teaching and Learning through Educational Data Mining and Learning Analytics" in 2012, pointing out that in the field of education, the specific application of big data is to analyze learning and mine educational data to improve the adaptive learning system and achieve personalized learning. Most foreign research focuses on educational evaluation, while in terms of student evaluation or the comprehensive quality evaluation of college students, the evaluation methods mainly focus on the evaluation of college students' academic ability and learning

effect.

Examining the research field of the comprehensive evaluation of college students from a global perspective, foreign research in this area appears relatively weak. Especially in the wave of the big data era, there is still a lack of research on how to specifically apply big data to learning behavior analysis and educational data mining to optimize the adaptive learning system and promote the development of personalized learning paths. Most foreign research focuses on the broad educational evaluation system, emphasizing the measurement of students' academic ability and learning achievements. In terms of constructing an evaluation system that comprehensively reflects the comprehensive quality of college students, the methods are mostly limited to the single dimension of academic performance and learning effect.

Zhang and Shi (2026) optimized the thermal insulation performance of transport packaging boxes using a response surface algorithm [1]. Zhu et al. (2026) employed proximal policy optimization for motion control of flexible-joint robotic arms in variable-station warehouse sorting [2]. Zhang (2026) proposed a hybrid LSTM-GARCH model integrating volatility factors from US financial markets [3]. Han et al. (2026) developed a dual-scale model with collaborative reasoning and multi-feature fusion for robust AI-generated image detection [4]. Shen et al. (2026) built a survival risk prediction model for colorectal cancer patients using graph convolutional networks and TCGA multi-omics data [5]. Tian et al. (2025) introduced a cross-attention multi-task learning approach for digital advertising recall [6]. Jiang et al. (2025) proposed UMH-ID, unifying multi-expert attention and hard contrastive learning for robust IoT intrusion detection [7]. Luo (2026) analyzed the integration of computer application technology and information management [8]. Ya (2025) studied EDA technology methodologies in digital circuit design [9]. Wang (2026) examined the application of computer science and technology in the big data context [10]. Wang (2025) developed a multi-scale feature-enhanced YOLOv8 for object detection in photovoltaic farm panoramic imagery [11]. Liu (2025) systematically evaluated deep learning architectures for plant image classification [12]. Gao (2025) researched the international intelligent operation model of "digitalization+greening" dual-track collaboration for industrial electrical enterprises [13]. Shen et al. (2025) applied the whale optimization algorithm to financial payment fraud detection [14]. Finally, Sun (2025) proposed adaptive interfaces for personalized user experience using a machine learning approach [15].

To sum up, according to the analysis of domestic and foreign research status, there is a common phenomenon that there are more theoretical studies and fewer operable studies on the comprehensive quality evaluation system of college students at present. Each university has its own characteristics, and the corresponding evaluation system should also be different, and the evaluation system should have sufficient pertinence. Moreover, with the advent of the big data era, every bit of college students' study and life will be tracked and recorded, and the comprehensive quality evaluation of college students should also progress with the changes of the times, provide feedback for the development of students' comprehensive qualities, help students' self-awareness and self-improvement, and at the same time provide a basis for universities to improve their education and teaching work. With its powerful data processing and learning capabilities, AI technology shows great potential in student evaluation. Through AI algorithms, the system can automatically analyze students' learning behaviors and learning processes, and provide personalized learning suggestions and guidance. At the same time, AI technology can also conduct personalized evaluations according to different evaluation indicators and weights, making the evaluation more fair and accurate.

## 3. DESIGN OF THE STUDENT COMPREHENSIVE EVALUATION SYSTEM

The student comprehensive quality evaluation system is a system designed to comprehensively evaluate students' academic achievements, various abilities and qualities. To achieve the goal of scientific evaluation, this project studies the student comprehensive quality evaluation system under the background of big data from aspects such as student behavior data collection, construction of evaluation index systems, and construction of comprehensive evaluation systems. The overall architecture of the system is shown in Figure 1.

### 3.1 College Student Behavior Data Collection

3.1.1 Sort out data standards and clarify the sources of students' behavioral data

The data sources of college students are characterized by wide distribution and diverse storage forms. Classified by business domain, the academic affairs system manages data such as students' academic records, grades, and teaching; the student affairs system manages data such as students' scholarships, grants, loans, and integrity penalties; the campus card system manages students' campus consumption data; the student union and youth

league system manages data of the second classroom, etc. Classified by storage form, structured data is stored in databases such as Oracle, MySQL, and SQL Server; unstructured data includes a large amount of text, images, audio, and video, etc.; file system data includes logs or log files, etc.; API interfaces include Web APIs and Web Service interfaces opened by third-party systems. To achieve the management of authoritative data sources, the following two aspects of work need to be carried out: First, policy documents should be issued to clarify the requirements for the types of data sources, storage forms, collection frequencies, etc. responsible for business departments; second, the data standard system should be sorted out to clarify the coding rules and standards for various types of data. For example, the student number coding is defined as 10 digits, consisting of the year (2 digits) + college code (2 digits) + major code (2 digits) + gender (1 digit) + process number (3 digits). The data standards should be iteratively updated regularly to ensure their guiding role.


**Figure 1:** Architecture Diagram of the Student Comprehensive Evaluation System

3.1.2 Establish a big data platform to collect students' behavioral data

Adopt a lakehouse data architecture to achieve data storage. Combining the characteristics of a data lake and a data warehouse, build a unified platform to store, manage, and analyze structured and unstructured data. This architecture includes data roles such as basic data, subject data, index data, and metadata. Basic data manages all structured and unstructured raw data, usually log files, JSON, CSV, pictures, videos, etc. in their original unprocessed or initial format and detail. Subject databases highly abstract data around 9 subject domains including organization, personnel, teaching, research, finance, assets, services, location, and public to support specific analysis tasks or business processes. Index data is a data structure created to improve query performance, which can help quickly locate and retrieve data. Metadata is data about data, describing the attributes and characteristics of data, such as the source, format, content, creation time, modification time, owner, access rights, etc. The lakehouse architecture provides a flexible and efficient data management and analysis platform by integrating the advantages of a data lake and a data warehouse, enabling universities to process various types of data on a unified platform and support complex data analysis and AI machine learning workloads.

### 3.2 Construction of Comprehensive Evaluation Index System

3.2.1 Data Governance: Building a Student Evaluation Index System

Data governance is the process of transforming data from a disordered state into an orderly aggregation. By formulating the metadata standards for school students' behaviors, building a student behavior metadata database, and setting up a metadata management platform, version management of metadata is achieved. Meanwhile, according to national and Ministry of Education standards, a reference data theme library is constructed to solve

the issues of data consistency and sharing for students.

Furthermore, by applying data mining methods, the implicit patterns and relationships in the data are explored, and students' performances are comprehensively analyzed from dimensions such as economy, academics, psychology, physiology, and behavior, so as to establish a more comprehensive and objective evaluation system. Economically, the return on investment in students' education is evaluated, that is, the relationship between students' investment in education (such as the costs of obtaining certificates, training, further education, etc.) and their future benefits (such as employment opportunities, salary levels, career development, etc.). Academically, students' knowledge mastery, learning skills, course grades, learning attitudes, self-learning abilities, academic integrity, critical thinking, application abilities, etc. are analyzed. Psychologically, students' motivation and interests, emotion regulation abilities, self-concepts, social abilities, adaptability, attention and concentration, creativity and imagination, etc. are analyzed. Physiologically, students' health conditions, physical development, sleep quality, eating habits, vision and hearing, fine motor skills, physiological adaptability, etc. are analyzed. Behaviorally, students' classroom behaviors, learning behaviors, social behaviors, self-management abilities, behavior habits, compliance with rules and disciplines, moral behaviors, etc. are analyzed.

In addition, through association analysis, the association rules between students' behaviors and performances can be identified. By setting appropriate support and confidence thresholds, the most meaningful association rules are filtered out, and these rules can serve as the basis for constructing evaluation indicators. Furthermore, clustering analysis will divide students into different groups according to their characteristics and performances. This unsupervised learning method can reveal the natural distribution and differences between student groups, helping school administrators understand which characteristics will classify students into a certain category. Combining the results of association analysis and clustering analysis, evaluation indicators for students' comprehensive qualities are constructed from multiple dimensions such as economy, academics, psychology, physiology, and behavior. This system can comprehensively consider students' economic situations, academic achievements, psychological states, physical health, and behavior habits, thus constructing a scientific evaluation system. This not only helps schools better understand students' needs but also enables the provision of personalized support and intervention measures for students.

3.2.2 Data Development, Constructing a Student Evaluation Algorithm Model

Based on the data index system, taking the theme data of the big data platform as data features, through the use of machine learning algorithms such as decision trees, random forests, and neural networks, a regression analysis is carried out on the student evaluation dimensions of moral education, intellectual education, physical education, aesthetic education, and labor education to construct a student evaluation algorithm model. The student evaluation algorithm model is described from two aspects: index description and data features, and evaluates the values of students' integrity, sense of responsibility, compassion, fairness, etc. and the students' cognitive abilities, learning abilities, knowledge mastery, and thinking skills from the five education dimensions respectively, and obtains data features from the corresponding indicators.

### 3.3 Construction of the Comprehensive Evaluation System

The student comprehensive evaluation system evaluates students' development from five dimensions: moral, intellectual, physical, aesthetic, and labor education, guides students' growth, teaching management, and job selection and employment, and realizes the automatic recommendation of students' learning suggestions, thus promoting the personalized development of students. The system functions mainly include single sign-on integration, student comprehensive evaluation analysis, student growth analysis, etc. The single sign-on integration is integrated with the unified authentication of the school online service hall to solve the problem of poor multi-account login experience for users; the student comprehensive evaluation analysis is displayed in a radar chart, with the dimension coordinates being the five-education evaluation dimensions. Student users can view their own evaluation models, and administrators can compare the evaluation models of multiple students at the same time. Students can query their own advantages and disadvantages in real time, playing the role of a guiding stick for student growth. The student growth analysis mainly combines the evaluation index system to explain the indicators of students' deficiencies and push suggestions in aspects such as student growth, education and teaching, and job selection and employment.

## 4. CONCLUSION

Based on the data of colleges and universities, this paper constructs a five-dimensional data index system for students, combines machine learning algorithms to build an evaluation model for the all-round development of students, and finally designs and develops a comprehensive student evaluation system to support students' personalized learning and the education and teaching management of colleges and universities. In the future, further research will be carried out on indicators and models such as students' psychological quality, social practice ability, teamwork and leadership, innovation and creativity, international vision and cross-cultural ability to improve the student evaluation system.

## REFERENCES

[1] Zhang, Y., & Shi, R. (2026). Optimal Design of Thermal Insulation Performance of Transport Packaging Box Based on Response Surface Algorithm. Innovation & Technology Advances, 4(1), 17–30. https://doi.org/10.61187/ita.v4i1.253

[2] Zhu, X., Pan, X., & Gao, X. (2026). Motion Control of Flexible-Joint Robotic Arms for Variable-Station Warehouse Sorting Based on Proximal Policy Optimization. Innovation & Technology Advances, 4(1), 1–16. https://doi.org/10.61187/ita.v4i1.296

[3] Zhang, X. (2026, March). A Hybrid LSTM-GARCH Model Integrating Volatility Factors from the US Financial Markets. In Proceedings of the 2026 International Conference on AI Decision-Making and Management (pp. 183-189).

[4] Han, J., Zhong, P. & Sun, L. Dual-scale model collaborative reasoning with multi-feature fusion for robust AI-generated image detection. Multimedia Systems 32, 363 (2026). https://doi.org/10.1007/s00530-026-02425-4

[5] Shen, Z., Lin, S., Wang, Y., Saunders, E., & Dai, Y. (2026, May). Survival Risk Prediction Model for Colorectal Cancer Patients Based on Graph Convolutional Network and TCGA Multi-Omics Data Integration. In 2026 7th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT) (pp. 370-373). IEEE.

[6] Q. Tian, D. Zou, Y. Han and X. Li, "A Business Intelligence Innovative Approach to Ad Recall: Cross-Attention Multi-Task Learning for Digital Advertising," 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shenzhen, China, 2025, pp. 1249-1253, doi: 10.1109/AINIT65432.2025.11035473.

[7] Jiang, K., Yuan, H., Zhang, T., Sun, L., Huang, S., & Xu, I. (2025). UMH-ID: Unifying Multi-Expert Attention and Hard Contrastive Learning for Robust IoT Intrusion Detection. Available at SSRN 7071219.

[8] Luo, R. (2026). The Integration Analysis of Computer Application Technology and Information Management. International Journal of Advance in Applied Science Research, 5(2), 27-30.

[9] Ya, L. (2025). EDA Technology in Digital Circuit Design: A Study on Application Methodologies. International Journal of Advance in Applied Science Research, 4(12), 6-10.

[10] Wang, J. (2026). Research on the Application of Computer Science and Technology in the Context of Big Data. International Journal of Advance in Applied Science Research, 5(1), 72-77.

[11] Wang, J. (2025). Multi-Scale Feature-Enhanced YOLOv8 for Object Detection in Photovoltaic Farm Panoramic Imagery. International Journal of Advance in Applied Science Research, 4(10), 7-11.

[12] Liu, Y. (2025). The Deep Learning Paradigm for Plant Image Classification: A Systematic Evaluation of Architectural Efficacy. International Journal of Advance in Applied Science Research, 4(8), 73-79.

[13] Gao, W. (2025, November). Research on the international intelligent operation model of "digitalization+ greening" dual-track collaboration of industrial electrical enterprises under intelligent drive. In Proceedings of the 2025 International Conference on Artificial Intelligence and Sustainable Development (pp. 157-162).

[14] Shen, Zepeng, et al. "Research on Application of Whale Optimization Algorithm in Financial Payment Fraud Detection." 2025 4th International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID). IEEE, 2025.

[15] Sun, L. (2025, November). Adaptive Interfaces for Personalized User Experience: A Machine Learning Approach. In Proceedings of the 2025 International Conference on Artificial Intelligence and Sustainable Development (pp. 457-462).