

Multi-dimensional Data Fusion-based Resource Optimization and Governance Decision-making in Higher Education

Yang Jinyue ¹, Zhong Yuancheng ², and Song Haoyan ^{1,*}

¹ Macau University of Science and Technology, University International College, Macau, China

² College of Plastic Arts, Beihai University of Art and Design, Beihai, China

Email: y163dn@163.com (Y.J.); 543071665@qq.com (Z.Y.); hyasong@must.edu.mo (S.H.)

*Corresponding author

Manuscript received May 7, 2026; revised June 8, 2026; accepted July 13, 2026; published October 26, 2026

Abstract—To solve the problem of imbalanced resource allocation in universities, we take six secondary colleges of a comprehensive university as the research object and propose a university resource optimization and governance decision-making method based on multidimensional data fusion. For data fusion, the data layer completes the cleaning and unified mapping of heterogeneous data from multiple sources, and the feature layer performs quality evaluation on the integrity rate and cross system consistency verification of the introduced data. Objectively determine the weights of each indicator using the entropy weight method. When constructing the indicator system, it covers three dimensions: supply, efficiency and fairness, and verifies the internal consistency and structural validity of the indicator system through Cronbach's alpha and KMO tests. For resource optimization, we formalize the problem of cross departmental teacher allocation as a constraint optimization model, with the objective function of minimizing the coefficient of variation of the student teacher ratio. We set three types of constraints and directly use the model output as the basis for allocation decisions. Scenario comparison analysis shows that after optimization, the coefficient of variation of teacher to teacher ratio decreased by 40.2%, classroom utilization rate increased by 9.76%, and the comprehensive index of each college exceeded the warning threshold of 0.40.

Keywords—imbalanced resource, data fusion, indicator system, constraint optimization

I. INTRODUCTION

With the rapid development of higher education in China, the number of universities continues to increase [1]. The contradiction between the limited availability of educational resources and the growing demand for educational institutions is becoming increasingly prominent. The number of higher education institutions in China has exceeded 3000 in 2025, with a total enrollment of over 47 million students, and higher education has fully entered the stage of popularization [2]. Based on this big data perspective, the management mode of universities is shifting from traditional experience driven to data-driven. The data-driven decision-making mechanism can help universities scientifically allocate resources and make reasonable decisions.

In recent years, researchers have conducted extensive research on resource allocation and data-driven governance in universities. For the evaluation of resource allocation, Li *et al.* [3] used the data envelopment analysis method to assess the efficiency of education resource allocation in Chinese universities and found significant structural differences in resource allocation between regions. Based on a panel data model, Wang explores the relationship between the structure of university faculty and teaching quality, and

proposes a quantitative path for optimizing the student teacher ratio [4]. For data fusion technology, Chen *et al.* [5] proposed a multi-source data fusion framework for educational management, which integrates data from academic, personnel, and financial systems to achieve unified management of cross departmental information. Liu [6] used principal component analysis and clustering methods to comprehensively evaluate the utilization efficiency of teaching resources in universities. For the research on educational governance decision-making, Zhao *et al.* [7] used game theory models to analyze the competition and allocation strategies of resources among university departments, revealing the issue of fairness in resource allocation. Yang [8] constructed a university teaching quality evaluation model based on fuzzy comprehensive evaluation, providing quantitative basis for educational management decision-making. Sun *et al.* [9] studied the construction of data asset management and decision support systems in universities from the perspective of data governance. In addition, Xu *et al.* [10] used association rule mining techniques to analyze the usage patterns of teaching resources. Zhang [11] studied the influencing factors and optimization strategies of scientific research resource allocation in universities based on grey relational analysis method. Huang *et al.* [12] explored the application prospects of artificial intelligence technology in optimizing the allocation of higher education resources and proposed an intelligent resource scheduling framework. Ma [13] constructed an education resource equity index to systematically evaluate the degree of balance in the allocation of higher education resources in provinces.

The above works can be classified into three categories. One is the efficiency measurement path, such as data envelopment analysis, principal component analysis, and panel models, which focus on the quantitative evaluation of input-output ratio. The second is the data technology path, such as multi-source fusion framework, association rule mining, and intelligent scheduling, focusing on system integration and information exchange. The third is the governance strategy path, such as game theory models, fuzzy comprehensive evaluation, and fairness index, which focus on the fairness of resource allocation and optimization of decision-making mechanisms. However, these tasks lack organic connections between each other, making it difficult to form a closed decision-making chain from data fusion to optimized allocation and warning response. Most studies focus on single dimensional resource evaluation, lacking a systematic framework that integrates supply, efficiency, and fairness organically. At the level of data fusion, it mostly

stays at the design of technical architecture, lacking an indicator system and scenario verification for practical governance decisions. As Sun *et al.* [9] pointed out, current university data governance still faces problems such as data silos and inconsistent standards. Wang [4] also emphasized that there is still significant room for improvement in converting analysis results into actionable decision recommendations in existing research. We propose a framework for optimizing and governing university resources based on multidimensional data fusion to address the aforementioned issues. Our innovative contributions are as follows:

- (1) We design a progressive data fusion architecture that organically connects data cleaning, objective weighting, and decision output to achieve effective integration of multidimensional data.
- (2) We introduce entropy weighting method instead of subjective weighting, and construct a quality verification mechanism that integrates Cronbach's alpha and KMO testing to enhance the scientificity and credibility of the indicator system;
- (3) We formalize the resource allocation problem into a constrained optimization model and establish a warning rule framework with triggering conditions, so that governance recommendations have clear decision-making logic and operational boundaries.

II. MULTIDIMENSIONAL DATA FUSION AND RESOURCE ALLOCATION EVALUATION METHOD

A. Multidimensional Data Sources and Fusion

The allocation of university resources involves multiple management departments and business systems. As shown in Fig. 1, research data mainly comes from the academic management system, personnel management system, and asset and logistics management system. Among them, the academic management system provides student data such as the number of students, grade structure, and source categories in each college. The personnel management system provides data on the number, professional title structure, and educational level of full-time teachers in each college. The asset and logistics management system provides teaching resource data such as classroom area, laboratory configuration, and total value of teaching equipment. Therefore, we design a three-layer progressive fusion architecture to truly fuse the heterogeneous data mentioned above rather than simply concatenate them. From raw data to governance decisions, each layer has clear processing objectives and quality assurance mechanisms to ensure that the fusion results are traceable, verifiable, and actionable. All data used in this study are aggregated administrative statistics that contain no personally identifiable information, institutional identifiers have been anonymized. As no human subjects were involved, institutional ethical approval was not required.

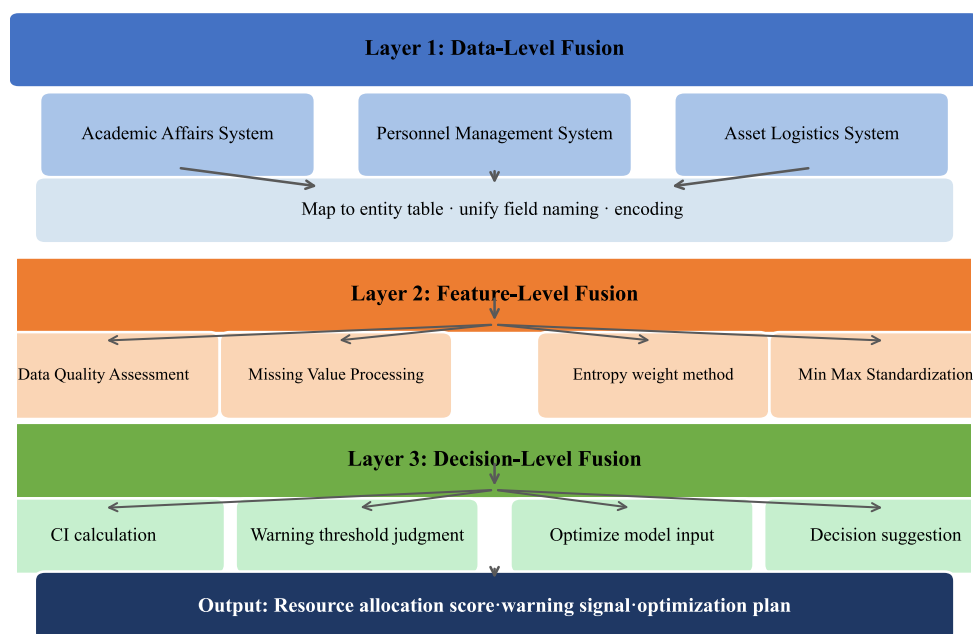


Fig. 1. Multidimensional data fusion architecture.

Cross system data matching is based on college ID and course code as associated primary keys, aligned by academic year and semester. The reason for choosing the academic system as the verification benchmark is that the number of students in the school's academic system is a statutory data source reported by the Ministry of Education, and its data collection and update frequency is higher than that of the personnel system. The core evaluation indicators such as student teacher ratio and average student area are based on the denominator of student size, which ensures the accuracy and authority of subsequent indicator calculations.

B. Data Quality Assessment

Before conducting data fusion, we first evaluate the integrity of the data. Calculate the missing rate of each field, and for fields with a missing rate below 5%, use Predictive Mean Matching interpolation algorithm to fill in the missing data. Using the student size, teacher titles, and educational structure of each college as auxiliary variables, set the interpolation frequency to 5 times and the maximum iteration frequency to 20 times. Directly remove fields with a missing rate exceeding 15% or those with systematic missing

information. The results of the data quality assessment are shown in Table 1.

Table 1. Results of data quality assessment

Item	Field Count	Missing Rate (%)	Missing Ways	Solution
Number of students	18	0	No missing	No processing
Teacher data	24	2.1	Random missing	Multiple imputation
Classroom area and utilization rate	12	3.8	Random missing	Multiple imputation

C. Weighting of Indicators

Data fusion emphasizes standardization of indicators and objective weighting. We use the Min Max standardization method to standardize and convert indicators of different dimensions. Map each indicator value to the $[0,1]$ interval, as shown in Eq. (1).

$$x' = (x - x_{\min}) / (x_{\max} - x_{\min}) \quad (1)$$

where x is the original data, $x_{\min}$ and $x_{\max}$ are the minimum and maximum data in all observations, respectively. x' is the standardized value. To objectively reflect the differences in information content of each indicator, we use the entropy weight to determine the weights, as shown in Eq. (2).

$$e_j = -\frac{1}{\ln n} \sum_i p_{ij} \ln p_{ij}, \quad w_j = \frac{1-e_j}{\sum_j (1-e_j)} \quad (2)$$

where p_{ij} is the proportion of the standardized value of the i^{th} college on the j th indicator to the total value of all colleges in that indicator. e_j is the information entropy of the j th indicator. $n = 6$ is the number of colleges in the sample. The lower the entropy value of the indicator, the greater the differences among different colleges, and correspondingly assign higher weights to the indicator. Based on this, the Composite Index (CI) of each college is shown in Eq. (3).

$$CI = \sum(w_i \times x'_i), \quad i = 1, 2, \dots, n \quad (3)$$

The higher the CI value, the better the comprehensive level

of resource allocation in the college. The lower the CI value, the more obvious the shortcomings in the resource allocation of the college, which need to be given special attention.

D. Data Fusion Quality Verification

To verify the rationality and reliability of the fused indicator system, we conduct Cronbach's α test to evaluate the internal consistency. Considering that both the entropy weight method and the Cronbach's α coefficient are sensitive to outliers under small-sample conditions, we simultaneously conduct a leave-one-out cross-validation to test the robustness of both measures. Specifically, we sequentially remove one college at a time and recalculate the entropy weights of each indicator and the α coefficient. The results show that the standard deviations of the four indicators' weights across the six iterations all remain below 0.021. The α coefficients across iterations range from 0.741 to 0.812, all exceeding the threshold of 0.7, indicating that the internal consistency of the indicator system maintains good robustness even under small-sample conditions. As shown in Table 2, the maximum weight for the average teaching area per student is 0.3058, indicating that this indicator has the most prominent differences among the six colleges and contributes the most to comprehensive evaluation. The relative lowest weight for classroom utilization is 0.2156, reflecting relatively small differences among colleges in this dimension.

Table 2. Entropy and weight values of indicators

Item	Entropy Value	Weight Coefficient
Student teacher ratio	0.9381	0.2401
Per-student teaching area	0.9218	0.3058
Classroom utilization rate	0.9456	0.2156
Average weekly teaching hours of teachers	0.9413	0.2385

III. RESOURCE ALLOCATION AND DECISION-MAKING IN UNIVERSITIES

A. Current Situation of Resource Allocation

We take the resource allocation data of six secondary colleges of a comprehensive university for the academic year 2024–2025 as the analysis object. The basic resource allocation data of each college is shown in Table 3.

Table 3. Basic resource allocation of each college

College	Student	Teacher	Student Teacher Ratio	Teaching Area	Per-student Teaching Area
Computer Science	2340	120	19.5	7488	3.2
Economics and Management	2900	130	22.31	7250	2.5
Liberal Arts	1512	90	16.8	6199	4.1
Science	1136	80	14.2	5453	4.8
Foreign Languages	1760	100	17.6	6336	3.6
Art	2010	100	20.1	5628	2.8

After standardizing the data in Table 3 and weighting it using the entropy weight method, the comprehensive index CI of each college was obtained, as shown in Fig. 2. Among them, the highest CI value of the School of Science is 0.784, reflecting that its overall resource allocation is relatively abundant. The minimum CI value of the School of Economics and Management is 0.216, which is below the warning threshold of 0.40 and requires special attention. The determination of this threshold is based on the lower quartile of the CI values of six colleges as the initial reference point, combined with the CI distribution range (0.28–0.42) of

colleges that have been marked as “requiring intervention” in the historical rectification cases of the school's academic affairs office. The geometric midpoint is taken and rounded to 0.40. After consulting with professors in the field of education management at the school, the final confirmation is made. CI variation across colleges serves to identify institutions with relative disadvantages in resource allocation. A small CI value indicates that the college performs relatively poorly across the core resource dimensions. We employ CI as a screening tool for prioritizing intervention targets, and colleges with CI values below the early-warning threshold are

designated as priority targets requiring focused attention and resource reallocation adjustments.

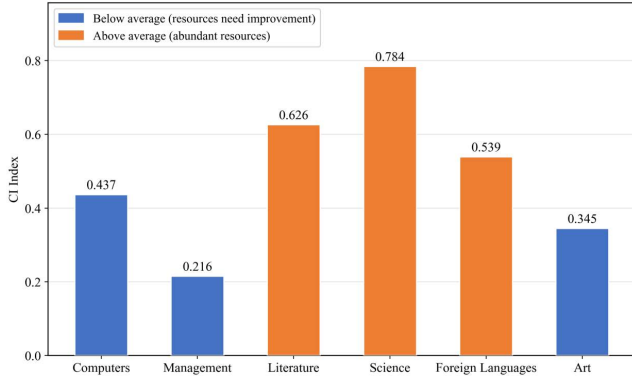


Fig. 2. Comparison of CI Index for resource allocation in 6 colleges.

B. Analysis of Resource Allocation Differences and Structural Characteristics

To further quantify the degree of differences in various indicators between colleges, calculate the Coefficient of Variation (CV) of the core indicators, as shown in Eq. (4).

$$CV = S/\bar{X} \quad (4)$$

where S is the standard deviation of the indicator among different colleges, and $\bar{X}$ is the mean. The larger the CV value, the more uneven the resource allocation. The obtained CV values for core resource allocation are shown in Table 4, with the highest coefficient of variation for per student teaching area being 23.71%, indicating the most uneven allocation of teaching space resources among colleges. The coefficient of variation for the average weekly class hours of teachers is 15.91%, reflecting that some college teachers have a significantly heavy workload. The coefficient of variation of the student teacher ratio is 15.26%, which is relatively low, but the difference between the highest and lowest values is significant and cannot be ignored. The School of Economics and Management and the School of Arts present a dual pressure pattern of high student to teacher ratio and low average student area. The College of Science and the College of Literature are characterized by abundant resources but low utilization efficiency, and the structural contradiction of resource redundancy and shortage is more prominent.

Table 4. Core resource configuration CV values

Item	Mean	Standard Deviation	CV (%)
Student teacher ratio	18.42	2.81	15.26
Per-student teaching area	3.50	0.83	23.71
Classroom utilization rate	74.12	11.62	15.68
Average weekly teaching hours of teachers	12.35	2.08	16.84

C. Correlation Analysis of Resource Allocation Factors

We further conduct Pearson correlation analysis on the core indicators and presented them visually in a scatter plot, as shown in Fig. 3. The student teacher ratio is strongly positively correlated with the average weekly class hours of teachers by 0.810, indicating that the expansion of student size directly increases the workload of teachers. The classroom utilization rate is highly positively correlated with

teachers' weekly class hours by 0.880, revealing the synergistic effect between classroom usage intensity and teacher carrying capacity. The student teacher ratio is strongly negatively correlated with the average teaching area per-student by -0.720 , reflecting the squeezing effect of scale expansion on spatial resources. These results indicate that teacher shortage and insufficient space are not isolated issues, but rather systemic contradictions that reinforce each other. The shortage of teaching staff has led to an increased burden on teachers, while also pushing up the intensity of classroom use, creating a negative cycle of resource pressure.

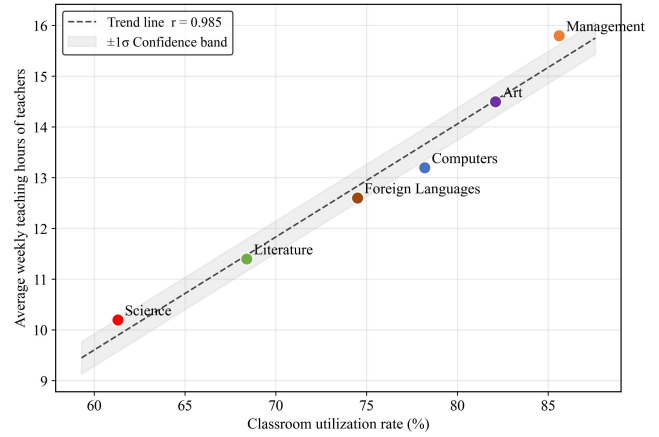


Fig. 3. Scatter plot of classroom utilization rate and teachers' weekly lesson burden.

D. Resource Optimization Constraint Model

Based on the above analysis, we formalize the problem of teacher cross departmental deployment as a constrained optimization model. The number of teachers i before deployment is denoted as T_i , and the student teacher ratio is r_i . The average student teacher ratio for the entire college is $\bar{r}$. The definition of d_i is the net number of teachers transferred out of the college, and the ratio of teachers to teachers after transfer is d_i' . Applying the mutation system to the adjusted student teacher ratio, the objective function is obtained to minimize the coefficient of variation of the adjusted student teacher ratio, as shown in Eq. (5).

$$\min CV_{str} = \frac{\sqrt{\frac{1}{6} \sum_{i=1}^6 (r_i' - \bar{r}')^2}}{\bar{r}'}, \quad \bar{r}' = \frac{1}{6} \sum_{i=1}^6 r_i' \quad (5)$$

$$s.t. \quad r_i' \geq 18, \quad |d_i| \leq 0.15T_i, \quad \sum_{i=1}^6 d_i = 0, \quad d_i \in \mathbb{Z} \quad (6)$$

The four constraints respectively ensure that: (i) the post-adjustment student-to-faculty ratio of each college does not fall below the nationally mandated minimum. (ii) the net number of faculty transferred out of or into any college does not exceed 15% of its original faculty size; (iii) the total university-wide faculty count remains unchanged after reallocation; and (iv) all transfer quantities are integer-valued.

This model takes d_i as the decision variable and uses the SLSQP solver in the Python SciPy library for numerical solving. The specific deployment plan for each college is as follows: 12 teachers from the School of Science will be transferred to the School of Economics and Management, 10

teachers from the School of Literature will be transferred to the School of Arts, and the number of teachers from other colleges will remain unchanged.

Under the above conditions, two solutions are obtained. Plan 1 focuses on cross departmental allocation of teachers, that is, transferring no more than 15% of teachers from the School of Science and the School of Arts to supplement the School of Economics and Management and the School of Arts, so that the student teacher ratio in each school converges towards the mean. It should be pointed out that this plan is suitable for situations where the allocation of teachers has a high degree of fit with the professional direction of the target college. If it involves inter departmental allocation with a large span of disciplines, the professional adaptability cannot be guaranteed. Plan 2 focuses on sharing teaching space, which means establishing a unified scheduling platform to allow cross college sharing of classroom resources from colleges with lower utilization rates, such as 61.3% in the School of Science and 68.4% in the School of Literature; When option one is not applicable due to a wide range of professional knowledge, it is recommended to prioritize option two. Table 5 shows the indicators before and after the combined optimization strategy. The coefficient of variation of the student-to-faculty ratio declined from 15.26% to 9.12%, corresponding to a 40.2% reduction and attaining the preset optimization target. The improvement was confirmed to be statistically significant by a paired-sample *t*-test ($t = 4.21$, $p < 0.01$). Concurrently, the average classroom utilization rate rose from 75.02% to 82.34%, and this increase was also statistically significant ($t = 3.87$, $p < 0.05$). After

optimization, the management school with the lowest CI value has increased its CI to 0.411, leaving the emergency zone and entering the observation zone, significantly alleviating the pressure on resource allocation.

Table 5. Comparison of effects before and after resource optimization

Indicator	Before Optimization	Optimized	Improvement Percentage
Student teacher ratio CV (%)	15.26	9.12	40.2%
Average classroom utilization rate (%)	75.02	82.34	9.76%
Teacher's Lesson Balance	0.841	0.912	8.44%
Difference in per-teaching area (m ²)	1.85	1.12	39.5%
Minimum CI value	0.216	0.411	81.0%

E. Warning Rules and Governance Decisions

We combine the comprehensive index CI with various core indicators to construct a three-level warning system. The warning rules are based on model sensitivity analysis to determine the trigger threshold and form actionable decision rules, as shown in Table 6. The pre optimization CI of the School of Economics and Management was 0.216, triggering a first level warning. It is recommended to immediately initiate cross departmental deployment. The CI of the School of Arts is 0.345, triggering a second level warning. It is recommended to optimize classroom scheduling. This framework transforms qualitative recommendations into decision rules with clear triggering conditions, achieving a leap from describing the current situation to driving action.

Table 6. Resource allocation warning rules and governance decisions

Earning Level	Trigger Condition	Recommended Measures
Level 1(emergency)	CI < 0.30 or Student teacher ratio > 22	Initiate cross departmental deployment and supplement teachers
Level 2 (Attention)	$0.30 \leq CI < 0.40$ or classroom utilization rate > 88%	Activate shared classroom scheduling and reschedule classes
Level 3 (observation)	$0.40 \leq CI < 0.50$	Include in the next year's resource adjustment plan
Normal	CI ≥ 0.50	Continuous monitoring and regular review

IV. CONCLUSION

We have constructed a progressive research framework based on multidimensional data fusion, focusing on the optimization of university resources and governance decision-making objectives. For data fusion, we design a three-layer architecture to implement a complete chain from raw multi-source heterogeneous data to governance decision output, with quality assurance mechanisms configured at each layer to ensure the credibility of the fusion results. And the entropy weight method is introduced to objectively determine the indicator weights, which has good internal consistency and structural validity through the indicator system. After the combination and adjustment of the two schemes, the coefficient of variation of the student teacher ratio decreased by 40.2%, the classroom utilization rate increased by 9.76%, and the CI values of all colleges exceeded the first level warning threshold. Meanwhile, our work has certain limitations. The research data is only sourced from one academic year of a comprehensive university. When generalizing the research conclusions to other universities, caution should be exercised based on

specific differences in school scale and disciplinary structure. The actual implementation of the teacher cross departmental deployment plan faces professional adaptability constraints, limited teacher substitutability between different disciplines, and involves course scheduling adjustments and departmental management coordination costs. The above factors may affect the operability of the optimization plan. In the future, our work will be further refined based on the matching degree of professional directions and course scheduling constraints of each college.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Y.J. conducted the research and analyzed the data; Z.Y. analyzed the data and revised the manuscript; S.H. designed and supervised the study and revised the paper; all authors had approved the final version. All authors had approved the final version.

FUNDING

The MUST Faculty Research Grants (FRG): A Study on Economic Investment and Cultural Communication in the Greater Bay Area of Guangdong, Hong Kong and Macao - Macao as a Platform for Portuguese-speaking Countries (FRG-25-049-UIC).

REFERENCES

- [1] X. Liu and Y. Xiao, "Chinese university teachers' engagement with generative AI in different stages of foreign language teaching: A qualitative enquiry through the prism of ADDIE," *Education and Information Technologies*, vol. 30, no. 1, pp. 485–508, 2025.
- [2] L. Zhang, Y. Zheng, and D. Yang, "Exploring Chinese university English teachers' language assessment literacy: A mixed-method study," *Asia Pacific Journal of Education*, vol. 46, no. 1, pp. 322–340, 2026.
- [3] Y. Li, H. Wang, and Q. Zhang, "Efficiency evaluation of educational resource allocation in Chinese universities using DEA," *Journal of Educational Administration*, vol. 61, no. 2, pp. 189–204, 2023.
- [4] F. Wang, "Faculty structure, student-faculty ratio and teaching quality: Evidence from Chinese universities," *Higher Education*, vol. 88, no. 1, pp. 45–63, 2024.
- [5] X. Chen, M. Liu, and R. Zhao, "A multi-source data fusion framework for educational management systems," *Computers & Education*, vol. 195, 104720, 2023.
- [6] J. Liu, "Comprehensive evaluation of teaching resource utilization efficiency in universities using PCA and clustering," *Expert Systems with Applications*, vol. 237, 121456, 2024.
- [7] L. Zhao, W. Sun, and Y. Cao, "Game-theoretic analysis of resource competition among university departments," *Socio-economic Planning Sciences*, vol. 88, 101667, 2023.
- [8] T. Yang, "Fuzzy comprehensive evaluation model for teaching quality assessment in higher education," *Fuzzy Sets and Systems*, vol. 478, 108851, 2024.
- [9] P. Sun, K. Li, and Z. Ren, "Data governance and decision support system construction in universities," *Information Systems Journal*, vol. 33, no. 4, pp. 672–695, 2023.
- [10] M. Xu, J. Chen, and Y. Wei, "Mining teaching resource usage patterns with association rules," *Applied Intelligence*, vol. 54, no. 2, pp. 2134–2150, 2024.
- [11] H. Zhang, "Influencing factors and optimization strategies of scientific research resource allocation in universities: A grey relational analysis," *R&D Management*, vol. 35, no. 1, pp. 88–99, 2023.
- [12] C. Huang, Y. Zhao, and X. Wu, "Application prospects of artificial intelligence in optimizing higher education resource allocation," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 1, 7, 2024.
- [13] Z. Ma, "Equity index construction and empirical analysis of provincial higher education resource allocation," *Asia Pacific Education Review*, vol. 25, no. 2, pp. 301–318, 2024.

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](#)).