An Enhanced Data Analysis Method for Teaching Evaluation Utilizing the Infomap Algorithm

Xiaoxu Xia[1], Min Gong[1], Kexi Liao[2], Chi Zhang[1,*], Jun Hong[2]
[1]*School of Sciences, Southwest Petroleum University, Chengdu, China*
[2]*Dean's Office, Southwest Petroleum University, Chengdu, China*

Abstract: With advancing educational informatization, universities now define evaluation indicators and collect extensive teaching data online. This study introduces an integrated analytical method for the evaluation of teaching that addresses persistent challenges of score inflation and limited discriminatory power in traditional methods. By combining the Continuous Golden Section Method (CGSM) with Infomap-based community detection, the method dynamically reconfigures scoring intervals using a golden-ratio-based algorithm (λ=0.618) to compress high-score regions and expand critical lower ranges, thereby resolving classification ambiguities caused by static thresholds. The network architecture employs bidirectional weighted connections derived from dual similarity criteria—k-nearest neighbors (k=6) and absolute distance thresholds (d=0.05)—to enhance clustering robustness in similarity-dense teacher populations through symmetric adjacency relationships. Multidimensional performance profiling is achieved via community-level mean analysis, enabling both comprehensive evaluation results of teachers and specific indicator diagnosis results. Unlike conventional weighted aggregation approaches, this method simultaneously tackles data distribution biases, network construction challenges in high-similarity environments, and holistic evaluation demands. Experimental validation using real-world university data demonstrates the model's viability, highlighting its ability to redefine rating ranges and improve distinctions among closely clustered teacher cohorts. The method advances educational evaluation by providing a mathematically rigorous, network-enhanced solution that prioritizes fairness and nuanced analysis, moving beyond oversimplified overall scores ranking to deliver actionable insights into specific teaching strengths and weaknesses.

Keywords: Teaching Evaluation; Internet Technologies; Complex Networks; Infomap; Golden Section

1. Introduction

With the increasing number of university students, the demand for excellent teachers is also increasing, and teachers' teaching ability has become the key factor affecting schools' teaching quality [1]. Education informatization has become mainstream, and the maturity of the network and corresponding technical facilities has provided the necessary platform support for online assessment questionnaires. Nowadays, most colleges and universities use online teaching evaluations in their assessment activities. Through scientific and reliable methods, the teaching-related department selects appropriate indicators to effectively reflect teachers' teaching quality. Then, the evaluation results are analyzed to obtain comprehensive scores or classifications, and finally, the related department provides feedback on those with low comprehensive scores or classifications as "unqualified" to the teachers themselves. Therefore , the choice of evaluation analysis method is particularly important, as it directly affects the accuracy of evaluation results and the improvement of teaching quality .While weighting-based methods dominate current teaching evaluations, they face critical limitations: e.g., the Analytic Hierarchy Process (AHP) relies on subjective expert assignments, introducing bias although nonlinear regression models have been proposed to optimize weights [2], these methods lack real-time adaptability and fail to pinpoint specific teaching issues (e.g., "insufficient classroom interaction"). To overcome these challenges, this paper proposes an unsupervised evaluation data analysis method based on complex network community detection, which offers:

Dynamic Adaptability: Automatically updating teacher communities via the Infomap algorithm without manual weight adjustments;

Granular Diagnostics: Identifying frequent weak indicators (e.g., A8 scores <0.6) within teacher communities, enabling targeted improvements.

The Entropy Value Method is used to determine weights and establish teaching evaluation models [3]. A Fuzzy Comprehensive Evaluation Model is utilized, which integrates the Analytic Hierarchy Process (AHP) with it. Subjective elements are minimized as much as possible through rigorous mathematical methods, and the weights of evaluation indicators are rationally determined [4]. The Entropy Weight TOPSIS weighted comprehensive evaluation model is constructed, utilizing this model to perform weighted teaching evaluations on various indicators [5]. The CRITIC technique is employed to derive the weight information based on Euclid distance and CSM technique under SVNSs [6]. Zhang explored the potential of edge computing and artificial intelligence to improve the quality evaluation of physical education (PE) teaching in ordinary colleges. The three-tier indicators were established, with the weights of the second and third-level indicators obtained through Q-learning. The indicator weight represents the degree of importance in conducting a top-to-bottom comparison of the evaluated items [7]. A multi-index evaluation method based on the fuzzy K-means clustering algorithm is designed [8]. This model constructs the online sports teaching effect evaluation framework, solves the index weights, and clusters the evaluation index parameters through the fuzzy K-means clustering method, thereby obtaining quantitative recursive results after constraining the parameters of convergence indices with nonlinear time series characteristics.

The above weight evaluation methods usually rely on available data and indicators, which may possess certain limitations. To ensure the accuracy of weight ratios, it is necessary to update the weights simultaneously with new data input. Therefore, the teaching evaluation is regarded as a classification problem. The Support Vector Machine (SVM) multiclassification method is introduced into the teaching level assessment task [9]. The application of a Weighted Naive Bayes Algorithm is offered to develop an intelligent teaching evaluation model [10]. In this study, an empirical algorithm serves as the basic framework to evaluate teaching quality, and the topic word distribution obtained by joint model training is used as the original knowledge. This approach enhances the traditional Naive Bayes Classifier by considering the correlations among features. Liu utilized the Apriori algorithm to optimize the teaching evaluation model, determining the indicators more correlated with teaching effectiveness. Then, an improved weighted Bayesian algorithm is proposed through incremental learning; it can provide the most likely class label for the new evaluation attribute value, which is the evaluation result [11]. The existing binary tree multi-classification algorithm is optimized, and a new classification algorithm is proposed [12]. Li proposes a simulation method for evaluating university teaching achievements [13], which is based on Deep Learning and an improved Vector Machine algorithm. The integrated neural network was adopted to evaluate the teaching quality of colleges and universities in [14], where the evaluation results of the RBF neural network, BP neural network, and echo state network were weighted to obtain the final results of teaching quality evaluation. Liu et al. proposed a multimedia teaching evaluation model based on deep convolutional neural networks and weighted Bayes. This model addresses the excessive subjectivity and randomness problems in teaching evaluation methods and uses the concept of class attribute correlation to evaluate the weight of each evaluation feature [15].

Nevertheless, these methods tend to focus on overall teacher performance, and the teacher's problem indicators are easily overshadowed by other good indicators. To solve this problem, Lin proposes a teacher evaluation method based on a multiple outlier detection approach [16]. However, it should be noted that since not all teachers are analyzed in this method, high accuracy in outlier detection is required. A performance evaluation model is proposed based on IoT and Bayesian network technology for operations research teaching. By utilizing the underlying prior probabilities obtained through field research, higher-level posterior probabilities are derived to analyze various performance standards that affect teachers' teaching [17].

Based on the above-mentioned problems, this paper proposes a method to measure not only the overall performance of teachers but also the performance of teachers' indicators. Current teaching evaluation systems confront two critical challenges: (1) Traditional weighting methods (e.g., entropy weighting [3], AHP [4]) and mainstream classification models (e.g., SVM [9], weighted Bayesian [11]) overly rely on holistic scores, homogenizing teachers' strengths/ weaknesses in specific indicators (e.g., "classroom interaction" or "content depth"), thereby failing to deliver precise improvement guidance; (2) Fixed grading thresholds (e.g., "≥90 for excellence") inadequately adapt to the prevalent high-score distribution, resulting in insufficient discriminability in upper score ranges and distorted evaluation outcomes.

To address these gaps, this study proposes two innovative solutions: First, a dynamic score interval partitioning mechanism based on the Golden Ratioiteratively compresses high-score intervals while expanding low-score intervals (Eq.1), enabling adaptive threshold adjustments to skewed data distributions. Second, a complex network model incorporating symmetric neighbor relationships defines node connections through dual constraints of k-nearest neighbors (k=6) and similarity thresholds (d=0.05), coupled with Infomap community detection to uncover group-level indicator patterns. Empirical validation demonstrates that our model identifies 32.9% of high-scoring teachers with specific indicator deficiencies (e.g., significantly lower A8 "post-class Q&A" scores in Community 14), offering administrators interpretable decision support that integrates global categorization and localized diagnostics.

2. Teaching Evaluation Data Analysis Model Based on Infomap

2.1 The Golden Section is Used for Grading

Due to students' respect for teachers and a perfunctory attitude in teaching evaluation, the teacher evaluation scores are consistently high. If the score range corresponding to the traditional grading system is used, the teachers' ratings may not be accurate. Therefore, this paper aims to propose a new method to reclassify the score range. It is assumed that the comprehensive score range of the evaluation data is [a, b]. The interval [a, b] can be divided into different grades using the Continuous Golden Section Method. If partitioned into n grades, the core idea is to first divide the entire range, then iteratively subdivide the intervals. Generally, poorly performing teachers are in the minority; thus, the Golden Ratio is developed (as shown in Figure 1), defining grades $T_1, T_2, \ldots, T_n$ with upper limits $t_1, t_2, \ldots, t_n$. The values of $t_1, t_2, \ldots, t_n$ can be obtained from Eq (1).

$$
\begin{aligned}
t_1 &= (1-0.618)^{n-1}b+\left[(1-0.618)^{n-2}+(1-0.618)^{n-3}+\ldots+(1-0.618)^{1}+(1-0.618)^{0}\right]0.618a,\\
t_2 &= (1-0.618)^{n-2}b+\left[(1-0.618)^{n-3}+(1-0.618)^{n-4}+\ldots+(1-0.618)^{1}+(1-0.618)^{0}\right]0.618a,\\
&\ldots\ldots,\\
t_i &= (1-0.618)^{n-i}b+\left[(1-0.618)^{n-i-1}+(1-0.618)^{n-i-2}+\ldots+(1-0.618)^{1}+(1-0.618)^{0}\right]0.618a,\\
&\ldots\ldots,\\
t_{n-2} &= (1-0.618)^{2}b+\left[(1-0.618)^{1}+(1-0.618)^{0}\right]0.618a,\\
t_{n-1} &= (1-0.618)b+\left[(1-0.618)^{0}\right]0.618a,\\
t_n &= b.
\end{aligned}
\tag{1}
$$

Figure 1. Continuous Golden Section Method

2.2 Complex Network

The basic elements of a complex network are nodes and edges. Nodes represent the components of a system, while edges represent interactions between components, with interaction strength quantified by edge weights. To construct a complex network for teacher evaluation systems, teachers are modeled as nodes, and relationships between them are analyzed. If a relationship exists, it is represented by a weighted edge.

Due to the inherent similarity in teacher evaluation data, this paper adopts a symmetric neighbor relationship [18], which effectively handles high-similarity datasets. During network construction, the following rules are applied:

Symmetric k-nearest neighbors (k-NN): For each node, its k-nearest neighbors (k=6) are identified. Bidirectional edges with maximum weight ($w_{ij}=1$) are assigned between nodes belonging to each other's k-nearest neighbors.

Similarity distance threshold: To mitigate limitations of fixed k-values, a distance threshold d=0.05 is introduced. Nodes with pairwise distances below d are strongly connected via bidirectional edges ($w_{ij}=1$), regardless of their k-NN membership.

The resulting complex network g integrates connections from both rules, ensuring robust community detection even in similarity-dense data regimes.

2.3 Infomap Algorithm

Infomap is a community algorithm used for

dividing complex networks into different communities. It performs random walks on a complex network and utilizes walk coding to represent information entropy [19]. Coding for complex networks includes assigning codes to each node as well as entry and exit codes for each community. A good community division ensures that before jumping to other community nodes, it traverses all nodes within its own community extensively, which significantly reduces code reuse for entering and exiting communities. Therefore, a short walk code length represents a good community division.
Building upon this idea further, access probability represents the proportion of occurrence for each node code within the community, while jump probability represents the proportion of occurrence for each community code.
In summary, we obtained complex network graph g through the aforementioned connections and divided g into different communities using the Infomap algorithm [20].

2.4 Mean Value Analysis

For each divided community, the overall mean analysis and indicators' mean analysis are conducted. This paper proposes that teachers within the same community exhibit similar performance; therefore, the community's mean score represents individual teachers' performance within it. The complete algorithm workflow is illustrated in Figure 2.
Here, X denotes the collected teacher data, K represents the number of nearest neighbors, and d is the distance threshold (customizable based on evaluation rigor requirements). The jumping probability τ is set to 0.15 [19].

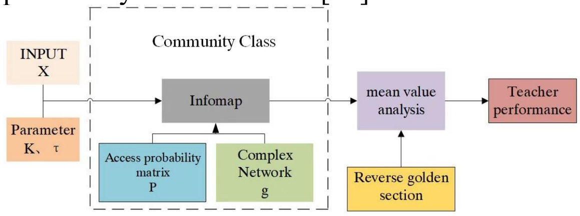

Figure 2. Algorithm Flow Chart

3. Experiment and Results

3.1 Data Preprocessing

The teaching evaluation data utilized in this paper are derived from the student evaluation of teaching system of a university during the autumn semester of the 2022 academic year. The original dataset comprises 12 fields, including the evaluation score, teachers' names, questionnaire name and number, course number, question number, course name, question name, college code of the reviewer, reviewer's college affiliation, and evaluator's name. In total, there are 127,389 data points. The question names represent the evaluation indicators and consist of 13 items denoted as A1 to A13 (see Table 1).

Table 1. Detailed Contents of Evaluation Indicators

Indicators	Contents
A1	Please provide an objective and fair comprehensive evaluation based on the overall teaching performance of the instructors for this course.
A2	The teaching content is closely aligned with the forefront of the discipline and industrial development.
A3	The teacher's availability for answering questions after class.
A4	The teacher's level of dedication and engagement in teaching this course.
A5	Incorporating elements of ideological and political education organically helps me establish correct views on life and values.
A6	Through studying this course, I have gained knowledge and skills.
A7	Serve as a role model, teach with dedication, manage the classroom effectively, and treat each student with respect and fairness.
A8	Aspects in need of improvement in teaching.
A9	Teaches in Mandarin with proficiency, highlights key points, and presents clear logic.
A10	Produces excellent multimedia courseware or maintains clear and standardized board work, resulting in effective teaching.
A11	The most commendable aspect of the teacher's instruction.
A12	Stimulates student interest and engages in interaction with students.
A13	The level of challenge in learning the course.

The field "questionnaire name" encompasses data on five major categories of teaching: classroom teaching, internship teaching, curriculum design, experimental teaching, and physical education skills. Data associated with classroom teaching categories were chosen for the research analysis, which comprised 113,828 data points. The four fields of the evaluation score, teacher number, question name, and evaluator name were then extracted as the follow-up evaluation data. Through the entropy method [3], we obtained the comprehensive evaluation of 400 teachers on each indicator (see Table 2).

3.2 Grade Division

Among teachers, 377 have an average indicator score ≥0.8, 16 fall within (0.7,0.8), and only 7 score at or below 0.7. This distribution aligns with the high-score tendency noted in our earlier analysis. In the Continuous Golden Section Method (CGSM), the upper limit b is set to the maximum observed mean (0.97). Since the minimum mean score (0.27) represents an extreme outlier, we exclude it from interval calculation to avoid distortion. Instead, a weighted lower limit a is computed by integrating two thresholds: 0.8 (applied to 377 teachers) and 0.27 (applied to 23 teachers). The final value of a is derived as Eq. (2):

Table 2. Comprehensive Scores of Some Teachers' Indicators

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13
T1	0.914	0.931	0.680	0.910	0.927	0.874	0.931	0.618	0.933	0.929	0.873	0.878	0.806
T2	0.936	0.936	0.841	0.936	0.936	0.907	0.936	0.500	0.936	0.936	0.861	0.936	0.773
T3	0.902	0.902	0.935	0.902	0.902	0.902	0.902	0.500	0.902	0.902	0.842	0.928	0.772
T4	0.960	0.960	1.000	0.960	0.960	0.960	0.960	0.506	0.960	0.960	1.000	0.960	0.802
T5	0.960	0.960	1.000	0.960	0.960	0.960	0.960	0.507	0.960	0.960	1.000	0.960	0.960
T6	0.960	0.960	1.000	0.960	0.960	0.960	0.960	0.500	0.960	0.960	1.000	0.960	0.960
T7	0.960	0.960	1.000	0.960	0.960	0.960	0.960	0.504	0.960	0.960	0.603	0.960	0.702
T8	0.897	0.891	0.768	0.891	0.897	0.897	0.897	0.598	0.897	0.897	0.791	0.891	0.862
T9	0.960	0.960	1.000	0.960	0.960	0.960	0.960	1.000	0.960	0.960	1.000	0.960	0.960
T10	0.960	0.960	1.000	0.960	0.960	0.960	0.960	1.000	0.960	0.960	1.000	0.960	0.90
T11	0.933	0.933	0.682	0.920	0.941	0.910	0.942	0.606	0.940	0.944	0.857	0.917	0.738
T12	0.942	0.938	0.706	0.922	0.920	0.900	0.946	0.556	0.952	0.943	0.814	0.907	0.865

The row index in table 2 represents each teacher with a serial number, and the column index represents the evaluation indicator. Each list element represents the score corresponding to different indicator for different teachers.

$$a = 0.8 * \frac{377}{400} + 0.27 * \frac{23}{400} \approx 0.77 \quad (2)$$

Consequently, [0.77,0.97] is the score range of the Reverse Golden Section. According to the above reverse golden section method, n = 3 is chosen, and the interval is split into three levels: $t_1 \approx 0.8$, $t_2 \approx 0.85$, $t_3 = 0.97$.

The quantitative analysis revealed a dominant proportion of instructors (classified as "good" with scores ≥0.85) demonstrating no significant need for pedagogical improvement, based on established performance benchmarks. A part of the data with qualified performance needs to be modestly improved. The score range is: 0.8≤pass<0.85.

A small number of teachers perform poorly, and they need to be greatly improved, the score range is: poor<0.8.

3.3 Teaching Evaluation Data Analysis

Complex Network Construction Rules:

Symmetric Neighbor Relationships: A bidirectional edge with weight $w_{ij}=1$ is established between teacher nodes i and j; Parameter Justification: Grid search validates that k=6 and d=0.05 yield modularity Q>0.7 (Figure 3), indicating significant community structure; Network Scale: 400 teacher nodes form a 400×400 adjacency matrix with 5,263 edges (density=3.3%). By selecting elements of adjacency matrix from which each row is not 0, the following series of network relations g are obtained (Table 3).

The complex network *g* is divided into different communities by the Infomap algorithm. To more objectively reflect the facts and determine the parameters, Modularity (Q) [21], the CH index, and the I index [22] were used to evaluate the community classification effect.

In the k-NN algorithm, selecting an appropriate k-value has a significant impact on model performance. When the k-value is small, the model becomes more susceptible to noise because it only considers a few neighboring points, which may lead to overfitting. The graph below shows the Q values for different k values (Figure 3).

The modularity (Q) is defined within the range [-0.5, 1], where higher values indicate a more significant division of community structure and stronger separation between network communities. However, in real-world network community detection, modularity values typically range from 0.3 to 0.7 [21].

Table 3. Teacher Complex System Network Matrix

Teacher Number	Interaction point effect (unidirectional)
0	0 29
1	1 206 219 283 332
2	2 130
3	3 42 120 342
4	4 5 17 23 44 55 98 190 199 220 231 232 235 240 258 265 271 287 338 367 381 394
5	4 5 17 23 44 55 98 190 199 220 231 232 235 240 258 265 271 287 338 367 381 394
6	6 238
7	7 12 141 172
8	8 9 36 39 40 78 79 86 89 126 132 143 166 169 185 187 223 244 251 259 272 276 278 297 309 316 335 352 362 365 371 373 379 385 390 393 396
9	8 9 36 39 40 78 79 86 89 126 132 143 166 169 185 187 223 244 251 259 272 276 278 297 309 316 335 352 362 365 371 373 379 385 390 393 396
10	10 74 139 155
11	11 15 77 100 321
12	7 12 90 141 188 226
13	13 43 195 391
14	14 142 164 294 324
15	11 15 65 77
16	16 150 178 202 320
17	4 5 17 23 44 55 98 190 199 220 231 232 235 240 258 265 271 287
⋮	⋮
397	93 284 304 307 397
398	63 77 296 398
399	47 250 399

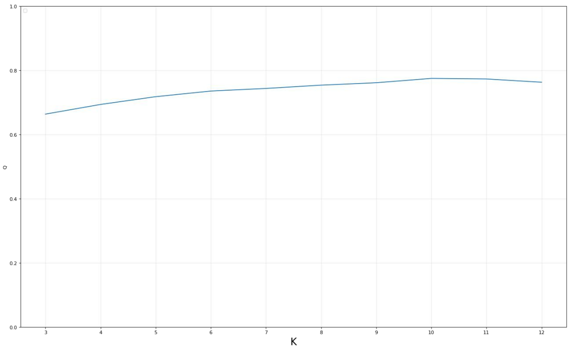

Figure 3. Q of Different K Values

As shown in Figure 3, although the modularity for K=9 (Q≈0.75) exceeds that for K=6 (Q≈0.73), both values approximate 0.7 with marginal difference. To rigorously determine optimal parameters, this study comprehensively evaluates the CH index, I index, and complementary metrics. Table 4 presents these metrics for K=6, 7, 8, and 9.

Table 4 demonstrates that the CH and I indices achieve optimal performance at K=6, showing significant superiority over other K values. Therefore, K=6 was selected as the optimal parameter. Applying the same methodology, we obtained d=0.05 through grid search optimization.

With these parameters, the model categorized 400 teachers into 103 communities. Each community contains teachers with homogeneous performance patterns across all indicators, enabling reliable representation of individual teacher profiles by community-level metrics.

To validate this, we selected sample 44 from community 4 for comparative analysis. As shown in Figure 4, the blue trajectory represents community 4's performance, while the orange trajectory depicts sample 44's scores. The close alignment between both trajectories confirms strong consistency in comprehensive and indicator-specific performance.

Table 4. The Index Value under Different K Values

K	Q	CH	I
6	0.735	910.2	10.3
7	0.745	825.8	8.9
8	0.753	817.9	8.45
9	0.759	790.3	8.1

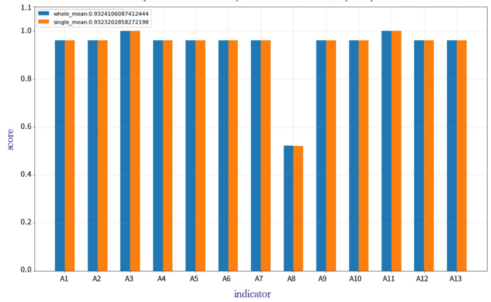

Figure 4. Comparison between the 4th Community and Sample 44

To further confirm our conclusions, another community was analyzed (see Figure 5); this figure compares Community 12 with samples 148, 218, 222, and 250 within that community. These samples exhibit fluctuations within a range closely aligned with the community's performance. Hence, the paper deems it viable to use community performance as a representation of teachers' performance within each community.

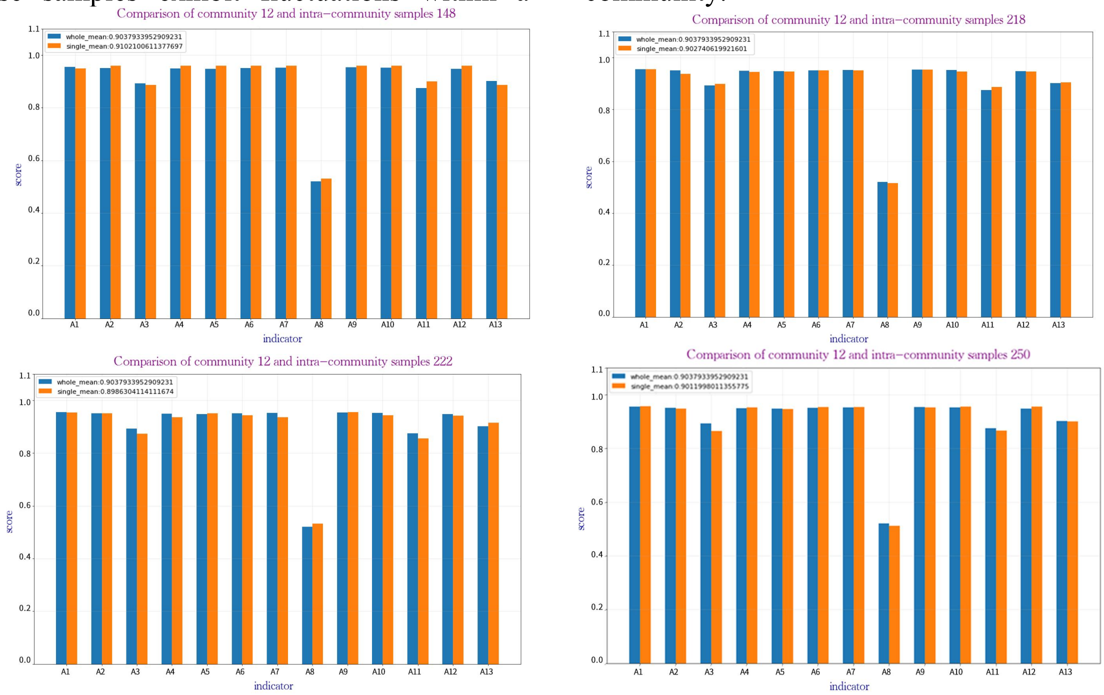

Figure 5. Comparison between the 12th Community and Sample 148, 218, 222, and 250

The mean scores for randomly selected communities (0, 1, 2, 3, 4, 5, 7, 8, 20, 30, 40, 50, 70, 80, and 102) are depicted in the left panel of Figure 6. Furthermore, random teacher samples from these communities are shown in the right panel of Figure 6. The two-line charts first reveal noticeable disparities between communities. Subsequently, each polyline demonstrates similar trends, indicating that the overall performance of a community mirrors the performance of individual teachers within it.

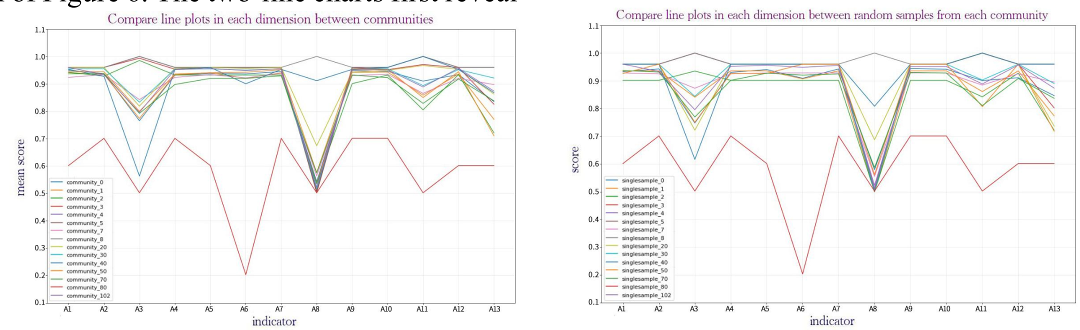

Figure 6. Comparison of Samples between Communities and within Communities

Case Study of Community 14:

Community Characteristics: All 10 teachers show "Good overall performance (mean=0.903) but pass-level A3 (Q&A Availability, mean=0.824)";

The hierarchical classification accuracy was 95% (124/130) with a 5% misclassification rate (6/130), complying with standard measurement error specifications; Improvement Suggestions: Targeted training programs for A3 to enhance after-class Q&A efficiency.

The data for teachers in the 14th community (Table5) are presented, which contains a total of 10 teachers. By calculating the average values for each indicator and the overall average (Table 5), the scores for each teacher can be represented by these averages. It can be concluded that in the 14th community, the performance on indicator A8 was poor; however, the performance on A3 was satisfactory, and the overall performance was good (Table 6). Therefore, the performance of teachers in the 14th community is shown in Table 6, where C represents the comprehensive score and "*" represents other indicators. Since the analysis focuses on indicators of poor performance, the tables highlight both the overall performance and areas requiring improvement. Other communities use similar methods to obtain teacher scores.

Table 5. Data of Teachers in the 14th Community

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13
T18	0.954	0.954	0.814	0.931	0.954	0.941	0.941	0.502	0.941	0.954	0.948	0.931	0.876
T35	0.960	0.960	0.844	0.960	0.960	0.960	0.960	0.544	0.960	0.960	0.946	0.960	0.855
T46	0.950	0.938	0.826	0.937	0.937	0.929	0.960	0.561	0.950	0.950	0.954	0.950	0.869
T91	0.960	0.960	0.858	0.960	0.960	0.946	0.960	0.547	0.914	0.960	0.952	0.960	0.884
T142	0.942	0.942	0.821	0.942	0.942	0.926	0.942	0.513	0.942	0.942	0.946	0.942	0.864
T170	0.960	0.956	0.791	0.960	0.960	0.960	0.960	0.509	0.956	0.960	0.981	0.956	0.820
T331	0.960	0.960	0.859	0.960	0.960	0.960	0.960	0.551	0.960	0.960	0.944	0.960	0.874
T347	0.956	0.956	0.820	0.956	0.956	0.956	0.956	0.515	0.956	0.956	0.955	0.956	0.876
T355	0.954	0.955	0.814	0.943	0.955	0.946	0.953	0.515	0.953	0.953	0.927	0.951	0.834
T389	0.960	0.960	0.798	0.960	0.960	0.960	0.960	0.501	0.960	0.960	0.959	0.960	0.855
mean	0.956	0.954	0.824	0.951	0.954	0.948	0.955	0.526	0.949	0.956	0.951	0.953	0.861

Table 6. Performance of Teachers in the 14th Community

	C	A3	A8	other
score	0.903	0.824	0.526	*
grade	good	pass	poor	good

4. Conclusion

The Infomap Algorithm is introduced in this paper as analytical method for the evaluation of teaching, allowing for the assessment of both overall teacher performance and performance on specific indicators. Initially, the limitations of current teaching evaluation analytical methods are examined. Following this, a new method proposed as Continuous Golden Section Method (CGSM) is utilized to redefine the grade range corresponding to traditional grades. Subsequently, by using a complex network and Infomap algorithm, teachers with similar performance levels are divided into a community. Then, comprehensive mean analysis and indicator mean analysis are conducted for each community, with the performance of the entire community considered representative of the performance of individual teacher samples within it. Through data experiment, the feasibility and effectiveness of this method in improving the teaching evaluation data analysis processes are demonstrated finally.

Acknowledgments

The authors are thankful to the esteemed reviewers for their valuable comments that significantly improved the quality of this article, and also extend gratitude for the project support. This work is supported by the projects "Research and Practice on the Reform of Blended Online and Offline Teaching in Advanced Mathematics Courses" [Grant No. 231101984035149] and "Research and Practice on the Reform of Classroom Teaching Quality Evaluation Methods Enabled by the New Generation Digital Technology." [Grant No.X2024JGYB58]

References

[1] Chen C. Y., Wang S. Y., Yang Y. F. A Study of the Correlation of the Improvement of Teaching Evaluation Scores Based on Student Performance Grades. International Journal of Higher Education, 2017, 6(2): 162-168.

[2] Zhao J. Z. Multiple regression analysis model and its application to teaching level prediction. Science and Technology Wind, 2018, (33): 45-46.

[3] Jiang H. F., Zhu W. J. Teaching evaluation model based on entropy method. Proceedings of the Journal of Hubei University of Technology, 2009, 24(4): 91-93.

[4] Ding J. L., Ye J. H. Application of Analytic Hierarchy Process and fuzzy comprehensive Evaluation in Classroom teaching quality evaluation. Proceedings of the Journal of Wuhan University (Humanities Edition), 2003, 56(2): 241-246.

[5] Yi W. H. Application of Weighted Comprehensive Evaluation Model in Teaching Quality Evaluation System of Computer Class Vocational Education. Applied Mathematics and Nonlinear Sciences, 2024, 9(1): 1-15.

[6] Yuan Y. Enhanced EDAS technique for colleges business English teaching quality evaluation based on Euclid distance and cosine similarity measure. Journal of Intelligent & Fuzzy Systems, 2024, 46(1):

75-89.

[7] Zhang J. B., Zhang C. Teaching quality monitoring and evaluation of physical education teaching in ordinary college based on edge computing optimization model. Journal of Supercomputing, 2023, 79(15): 16559-16579.

[8] He J. J., Ye M. L. Evaluation method of online teaching effect for physical education based on fuzzy k-means clustering. Proceedings of Jilin University (Information Science Edition), 2023, 40(2): 301-306.

[9] Sheng W. X., Wang H., Dong X. R., Xie G. H. Research on support vector Machine-based teaching level evaluation model in universities. Proceedings of Computer Knowledge and Technology, 2019, (13): 3165-3168.

[10] Liu L. Smart teaching evaluation model using weighted naive bayes algorithm. Journal of Intelligent & Fuzzy Systems, 2020, 40(2): 1-11.

[11] Gu Y. R. Exploring the application of teaching evaluation models incorporating association rules and weighted naive Bayesian algorithms. Intelligent Systems with Applications, 2023, 20: 200297.

[12] Lou M. S. Evaluation of College English Teaching Quality Based on Improved T-SVM Algorithm. Computational Intelligence and Neuroscience, 2022, 2022: 2974813. 10.1155

[13] Li C. Simulation of University Teaching Achievement Evaluation Based on Deep Learning and Improved Vector Machine Algorithm.". Applied Artificial Intelligence, 2023, 37(1): 2195221.

[14] Wang L., Zhang H. J. Research on teaching quality evaluation system with integrated neural network. Proceedings of Modern Electronic Technology, 2021, 44(3): 6-73.

[15] Liu T., Ning L. Deep convolutional neural network and weighted Bayesian model for evaluation of college foreign language multimedia teaching. Proceedings of Wireless Communications and Mobile Computing, 2021, 2021(3): 1-7.

[16] Lin X. Q., "Research on teaching evaluation methods based on multivariate outlier detection," East China jiaotong university.

[17] Kong L. Performance evaluation model for operation research teaching based on IoT and Bayesian network technology. Soft Computing, 2024, 28: 3613-3631.

[18] Ke W. J., Wei J. G., Xiong N. X., Hou Q. Z. GSS: A group similarity system based on unsupervised outlier detection for big data computing.". Proceedings of Information Sciences 2023, 619: 1-15.

[19] Rosvall M., Axelsson D., Bergstrom C.T. The map equation. European Physical Journal Special Topics, 2010, 178: 13-23.

[20] Rosvall M., Ergstrom C. T. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences, 2008, 105(4): 1118-1123.

[21] Li H., Chen F. C., Zhang J. P., Wu Z., Li S. M., Huang R. Y. Community discovery algorithms in complex networks. Application Research of Computers, 2021, volume 38 (6): 1611-1618.

[22] Liu Y. C., Gao X. D., Guo H. W., Wu S. Combination evaluation method of cluster validity. Computer Engineering and Applications (19): 15-17.

