

# Knowledge-Based English Teaching Quality Assessment Methods: Using Data Mining Techniques to Identify Key Factors and Improve Teaching Strategies

**Yun Cao**

Institute of Foreign Languages, Harbin University of Commerce, Harbin 150028, Heilongjiang, China

Corresponding author: Yun Cao (e-mail: helenyun20241224@163.com)

**ABSTRACT** Effective evaluation of complex learning systems requires reliable information acquisition, multidimensional data analysis, and adaptive feedback mechanisms. This study proposes a data-driven assessment framework for teaching-quality evaluation based on knowledge-oriented learning analytics and data-mining techniques. A multidimensional data acquisition architecture is developed to integrate classroom interaction records, online learning behaviors, assessment outcomes, and learner-affective information into a unified information-processing workflow. Multiple analytical modules, including cluster analysis, association-rule mining, sequence-pattern discovery, and predictive modeling, are employed to identify latent relationships between learning behaviors and educational outcomes. To improve evaluation reliability, a dynamic comprehensive indicator model is established for continuous quality monitoring and anomaly detection. Furthermore, an intelligent feedback mechanism is introduced to generate diagnostic reports, support adaptive strategy adjustment, and establish a closed-loop optimization process. The proposed framework enables systematic information extraction, pattern recognition, and decision support from heterogeneous educational datasets. By integrating data acquisition, intelligent analytics, and adaptive feedback, the framework provides an engineering-oriented methodology for information processing, behavior modeling, and dynamic performance assessment in complex data-intensive environments.

**INDEX TERMS** data mining, information acquisition, pattern recognition, adaptive feedback systems

## I. INTRODUCTION

ENGLISH teaching quality evaluation is a crucial means of ensuring teaching quality and promoting the reform of specialized English curricula designed for engineering professionals. Establishing robust assessment standards for technical communication ensures that students master the precise terminology and international standards required for advanced manufacturing and global fiber research. With the development of educational informatization, the teaching process generates a large amount of data, providing rich data resources for teaching quality evaluation. Data mining technology enables the analysis of extensive datasets to identify latent patterns, offering an empirical framework to support teaching quality assessment. However, current English teaching quality evaluation still suffers from problems such as insufficient data utilization, a single evaluation method, and inadequate application of results. Therefore, exploring English teaching quality evaluation methods based on data mining technology to identify key factors affecting the proficiency of engineering students in technical terminology has significant theoretical and practical implications. Refining these teaching strategies ensures that future profes-

sionals can accurately communicate complex manufacturing processes and fiber innovations within the global industrial landscape.

## II. THE SIGNIFICANCE OF DATA MINING-BASED EVALUATION OF ENGLISH TEACHING QUALITY

### A. ACHIEVING ACCURATE DIAGNOSIS AND ASSESSMENT OF ENGLISH TEACHING QUALITY

Data mining technology possesses the potential to deeply mine and analyze various types of data generated during the teaching process. By integrating fragmented data points, it offers a framework to more objectively reflect the real situation of English teaching compared to traditional methods, provided that the underlying data collection infrastructure is sufficiently robust [1]. Traditional evaluation methods rely on final exam scores and teachers' subjective judgments, and the evaluation dimensions are relatively simple, which cannot reflect the problems existing in teaching well. Data mining technology can integrate information from various aspects such as students' classroom performance data, homework completion status, online learning behavior, and test scores. Compared to standard linear regression, which

often fails to capture non-linear relationships and complex interactions in pedagogical settings, algorithms such as cluster analysis and association rule mining can identify multi-dimensional patterns and latent student groupings without predefining dependent and independent variables [2].

### **B. PROMOTING THE SCIENTIFIC AND PERSONALIZED NATURE OF ENGLISH TEACHING DECISIONS**

Data mining technology can discover the inherent connections between teaching elements and provide a scientific basis for teaching decisions. By mining and analyzing historical teaching data, we can discover the correlation between various factors such as teaching methods, teaching resources, and student characteristics and learning outcomes, and find the main factors affecting teaching quality [3]. These findings can provide teachers with a basis for the actual effects of different teaching strategies, enabling them to know which teaching behaviors are most helpful to students' academic progress, and thus make corresponding adjustments to the teaching content and pace [4].

### **C. PROMOTING THE INNOVATION AND OPTIMIZATION OF ENGLISH TEACHING MODELS**

Data mining technology provides new perspectives and tools for the innovation of English teaching models. Traditional teaching models often adhere to existing teaching processes and methods, lacking dynamic monitoring and timely adjustment of teaching effects. Data mining technology can collect and analyze teaching process data in near-real-time or at regular intervals, providing iterative feedback for the periodic improvement of teaching models [5]. Teachers can rely on data analysis to continuously adjust and improve their teaching designs, thereby identifying optimized pedagogical organizational methods and enhancing the data-driven precision of the instructional process [6]. Through the above analysis, we have derived the English teaching quality evaluation implementation process established in this paper, the specific composition of which is shown in Figure 1, consisting of four steps: data collection, mining analysis, quality evaluation, and feedback improvement.

## **III. PROBLEMS IN THE CURRENT EVALUATION OF ENGLISH TEACHING QUALITY**

### **A. THE EVALUATION DATA COLLECTION WAS INCOMPLETE AND LACKED IN-DEPTH ANALYSIS**

The current data collection in the evaluation of English teaching quality is often characterized by being one-sided and superficial, which inherently limits the accuracy of any subsequent data mining outputs. Many schools only use the final and midterm exam scores as the evaluation basis, and lack the collection of process data such as students' classroom participation, homework completion quality, and self-learning behavior. Even if some schools have established online learning platforms, they only record superficial data such as the number of logins and the duration of video viewing, and do not effectively track students' deep learning

behavior. The data collection time span is small and cannot form a complete learning trajectory record [7]. The more prominent problem is that most of the collected data is still at the level of simple statistical summary, without in-depth mining and correlation analysis, which makes the evaluation results unable to truly reflect the teaching situation [8].

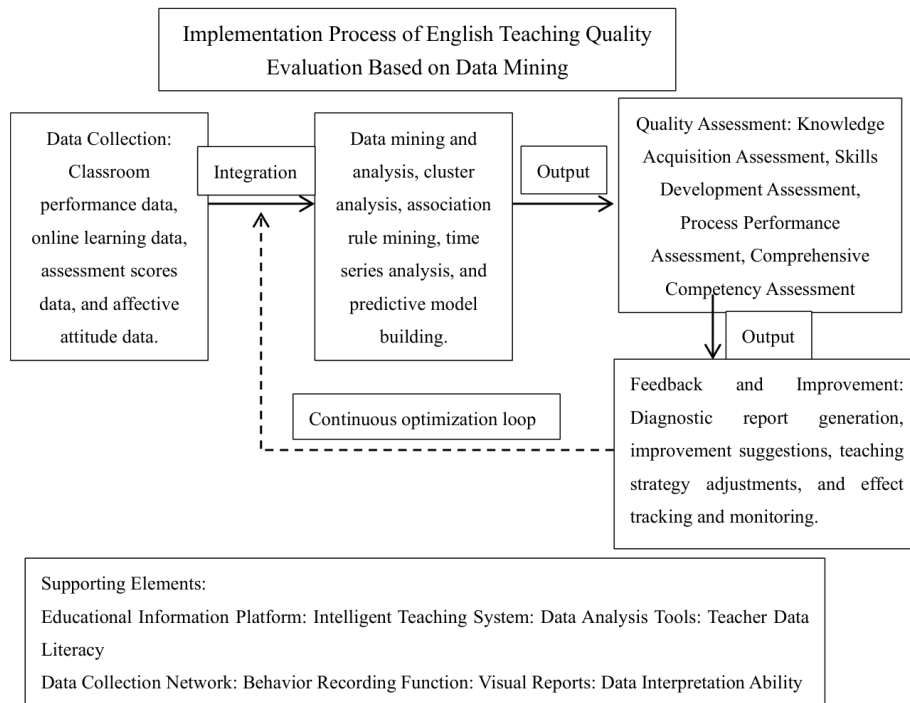
### **B. THE EVALUATION INDEX SYSTEM IS TOO SIMPLISTIC TO REFLECT THE OVERALL TEACHING SITUATION**

The current English teaching quality evaluation index system has the problems of being too singular and one-sided, and cannot fully and objectively reflect the true situation of teaching quality. Most evaluation systems rely too much on summative evaluation standards based on exam scores, and take students' test scores as the main or even the only evaluation standard for measuring teaching effectiveness. These assessment metrics often ignore the comprehensive and practical nature of language acquisition. Consequently, they fail to reflect balanced development across listening, speaking, reading, and writing skills. Furthermore, they provide little insight into students' functional language application or intercultural communication literacy [9]. The evaluation indicators lack attention to the teaching process, the evaluation standards are relatively fixed, and cannot adapt to the needs of students of different grades and different types of growth. They cannot accurately diagnose the problems existing in teaching, nor can they provide a clear direction for teaching improvement [10].

### **C. INSUFFICIENT APPLICATION OF EVALUATION RESULTS FAILED TO EFFECTIVELY GUIDE TEACHING IMPROVEMENT**

The application of the results of English teaching quality evaluation generally suffers from the problem of emphasizing evaluation over application, and an effective closed loop cannot be formed between evaluation and teaching improvement. Most schools conduct teaching quality evaluations regularly, but the evaluation results are often used as the basis for teacher assessment or presented in the form of simple grade reports without detailed interpretation and analysis. Teachers cannot obtain specific and actionable improvement suggestions from the evaluation results and do not know which aspects to start from to improve teaching quality [11]. The feedback of evaluation results is not timely enough, and is generally announced at the end of the academic year, thus losing the best time for teaching adjustments. More importantly, the evaluation results are not linked to specific teaching strategies, making the evaluation a mere formality and unable to truly promote teaching improvement and enhance teaching quality [12].

## **IV. DATA MINING-BASED STRATEGIES FOR IMPROVING THE QUALITY OF ENGLISH TEACHING**



**FIGURE 1.** Flowchart of English Teaching Quality Evaluation Based on Data Mining

#### A. CONSTRUCTING A MULTI-DIMENSIONAL DATA ACQUISITION SYSTEM AND APPLYING DATA MINING ALGORITHMS FOR IN-DEPTH ANALYSIS

Creating a comprehensive and systematic multidimensional data collection system is the essential prerequisite for bridging the aforementioned gap and implementing data-driven teaching quality evaluation. Only by addressing the current "data silos" and "superficiality" can the theoretical advantages of data mining be translated into practical pedagogical insights. Schools should make full use of the education information platform to establish a data collection network for the entire learning process of students and systematically collect information on various aspects such as students' classroom performance data, online learning behavior data, homework test data, and language skill development data [13]. In the classroom teaching process, intelligent teaching systems are used to record behavioral data such as the number of times students speak, the frequency of interaction, and the degree of attention. Online learning platforms need to track in-depth learning data such as the students' learning process, the use of each resource, and the quality of each exercise. In addition, attention should be paid to students' emotional attitude data and self-evaluation data to understand the students' learning situation in an all-round way [14]. In order to ensure the comprehensiveness and systematicness of data collection, this paper designs a data collection framework with four main aspects as the content (see Table 1).

**TABLE 1.** Multidimensional Data Acquisition Framework for English Teaching Based on Data Mining

| Data Dimensions      | Content collected                                              | Data collection tools/methods                      | Application Algorithm   |
|----------------------|----------------------------------------------------------------|----------------------------------------------------|-------------------------|
| Classroom behavior   | Speaking frequency, interactive participation, attention state | Intelligent teaching system, classroom observation | Cluster analysis        |
| Online learning      | Learning path, resource usage, and practice quality            | Online learning platforms, log tracking            | Sequence pattern mining |
| Academic performance | Homework grades, test results, skill progress                  | Assessment system, electronic assignment           | Association rule mining |
| Emotional attitude   | Learning interest, self-evaluation, and motivation level       | Questionnaire survey, learning log                 | Text mining             |

Note: The selection of these algorithms is based on their ability to handle categorical and sparse behavioral data common in specific classrooms. For instance, Sequence Pattern Mining is employed to capture the temporal progression of learning paths, a dimension often overlooked by traditional regression models. Based on data collection, cluster analysis algorithms are used to classify students' learning situations and identify different student groups. Association rule mining algorithms are used to find the relationship between teaching elements and learning outcomes, and to identify the factors that affect students' learning outcomes. Time series analysis is used to track students' learning progress and predict the development trend of learning. By using these data mining techniques, useful information and knowledge can be found from teaching data. While these methods are often associated with "big data," specific im-

plementations such as K-means for small-scale clustering or Apriori for sparse transactional logs are highly compatible with classroom-level datasets. These algorithms effectively handle the inherent sparsity of student interaction data by focusing on frequent patterns and local similarities rather than global statistical significance, providing a robust basis for the evaluation of teaching quality [15].

### B. ESTABLISH A DYNAMIC COMPREHENSIVE EVALUATION INDEX MODEL TO ACHIEVE COMPREHENSIVE QUALITY MONITORING

To ensure the effectiveness of teaching quality evaluation, a scientific and reasonable comprehensive evaluation index system must be established. Based on the core competencies requirements of the English language subject and the goals of students' all-round development, a three-dimensional evaluation index system should be created, encompassing knowledge acquisition, skills development, learning process, and affective attitudes. The knowledge dimension examines students' understanding and mastery of language knowledge; the skills dimension examines the development level of each language skill (listening, speaking, reading, and writing); the process dimension examines students' learning engagement, self-learning ability, and collaborative learning performance; and the affective dimension examines students' learning interest, learning motivation, and intercultural awareness. The indicators of each dimension should be interconnected, forming an organic whole. In accordance with the requirements of the core competencies of the English language subject, this paper establishes a comprehensive evaluation system including 4 primary indicators and 12 secondary indicators (see Table 2).

**TABLE 2.** Dynamic Comprehensive Evaluation Index System for English Teaching Quality

| Primary indicators | Secondary indicators    | Evaluation criteria                            |
|--------------------|-------------------------|------------------------------------------------|
| Knowledge mastery  | Vocabulary              | Vocabulary recognition and application ability |
|                    | Grammar                 | Grammar rules mastery                          |
|                    | Cultural knowledge      | Cross-cultural cognition level                 |
| Skills Development | Hearing ability         | Information Acquisition and Understanding      |
|                    | Spoken expression       | Fluency and accuracy                           |
|                    | Reading comprehension   | Textual Analysis and Criticism                 |
| Learning process   | Writing ability         | Discourse Construction and Expression          |
|                    | Classroom participation | Interaction frequency and quality              |
|                    | Self-directed learning  | Learning Plan and Execution                    |
| Emotional attitude | Collaborative learning  | Teamwork performance                           |
|                    | Learning interest       | Initiative and proactivity                     |
|                    | Learning Motivation     | Intrinsic drive                                |

The evaluation index system should have a dynamic adjustment mechanism, allowing for flexible adjustments based on different learning stages and teaching objectives. The analytic hierarchy process (AHP) is used to deter-

mine the weights of each level of indicators, ensuring the scientific and reasonable nature of the evaluation. Data mining techniques can be used to dynamically monitor the changing trends of each indicator, enabling timely detection of fluctuations or anomalies in teaching quality. Once the early warning mechanism is established, the system will automatically issue an alarm if a certain indicator shows a significant decline, reminding teachers to pay attention and take intervention measures, thereby achieving the goal of systematic monitoring and dynamic control of teaching quality within feasible processing windows. Based on the three core issues in English teaching quality evaluation, this paper proposes an improvement strategy based on data mining techniques. Addressing the three core issues currently existing in English teaching quality evaluation, this study proposes a systematic improvement strategy based on data mining techniques (see Table 3).

**TABLE 3.** Problem Diagnosis and Data-Driven Improvement Strategies for English Teaching Quality Evaluation

| Core issues                         | Main manifestations                                                                                                                       | Improvement Strategy                                                                                                                                                                                                             | Expected Results                                                                                                                           |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------|
| Incomplete data collection          | Over-reliance on exam scores, lack of process data; primarily superficial records, insufficient in-depth tracking.                        | Construct a multi-dimensional data collection system to integrate data from multiple sources, including classroom, online, and homework data; and use algorithms such as clustering and association rules for in-depth analysis. | Comprehensively reflects the true state of teaching and learning, and accurately identifies learning characteristics and problems.         |
| Single evaluation indicator         | The assessment is primarily summative, neglecting comprehensive qualities; the standards are rigid and fail to adapt to diverse needs.    | Establish a dynamic comprehensive evaluation model covering four dimensions: knowledge, skills, process, and emotion; set up an early warning mechanism to monitor quality fluctuations in real time.                            | Objectively evaluate the overall teaching situation and identify quality abnormalities following the completion of data processing cycles. |
| Insufficient application of results | Evaluation and improvement are disconnected, lacking specific guidance; feedback is delayed, missing the best opportunity for adjustment. | Design an intelligent feedback mechanism to generate personalized diagnostic reports; provide targeted improvement suggestions and establish an effect tracking system.                                                          | Forming an evaluation-improvement closed loop to continuously enhance teaching quality.                                                    |

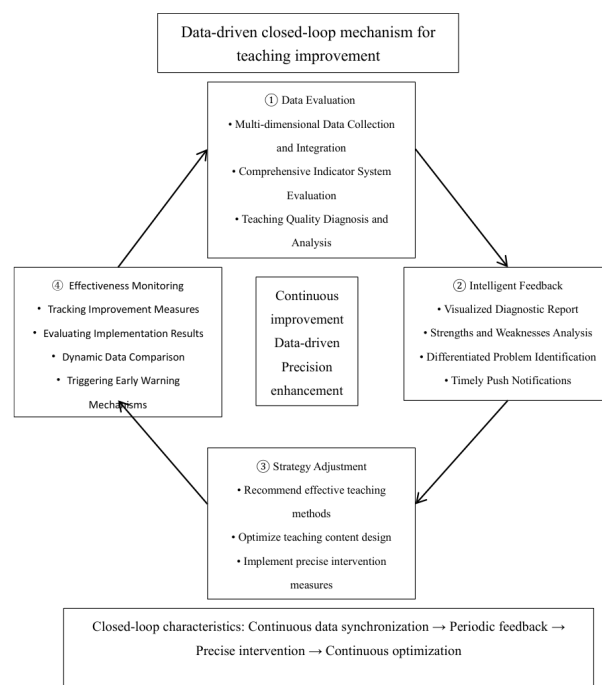


FIGURE 2. Data-driven teaching improvement closed-loop mechanism diagram

### C. DESIGNING A DATA-DRIVEN INSTRUCTIONAL FEEDBACK MECHANISM TO PROMOTE PRECISE IMPROVEMENT

To ensure the effective transformation of evaluation results, this study established a data-driven closed-loop mechanism for teaching improvement (see Figure 2), thereby ensuring that the evaluation, feedback, improvement, and monitoring processes are closely linked and continuously cyclical.

Establishing a timely and effective teaching feedback mechanism is a crucial step in evaluating and promoting teaching improvement. An intelligent teaching quality analysis system should be developed to automatically analyze and process evaluation data, generating intuitive and easy-to-understand visual analysis reports. The system should provide each teacher with a personalized teaching diagnostic report, showing their strengths and weaknesses and providing specific directions for improvement. The report should include an analysis of the overall student learning situation, the differences between different student groups, an assessment of the achievement of teaching objectives, and an analysis of the effectiveness of teaching methods. More importantly, the system should provide teachers with specific teaching improvement suggestions based on data mining results. By analyzing successful teaching cases in similar situations, effective teaching strategies and methods should be provided to teachers. Different teaching suggestions should be offered based on different student categories to implement precise teaching. The feedback mechanism must be efficiently scheduled, pushing interim analysis reports to teachers following a necessary period for data cleaning and algorithmic processing after each teaching

unit. This ensures that adjustments to teaching methods are based on verified, high-quality data outputs rather than raw, unrefined metrics. A teaching improvement effect tracking and evaluation mechanism should be established to track and analyze the implementation of teaching improvement measures, constructing a closed-loop process for improving English teaching quality and continuously improving the level of English teaching education.

### V. CONCLUSION

Data mining technology has brought new perspectives and methods to the evaluation of English teaching quality by accurately analyzing student mastery of specialized terminology used in machinery and fiber science, effectively addressing the shortcomings of traditional evaluation methods. This analytical approach enables educators to refine language training to meet the precise communicative standards required for international technical cooperation in the manufacturing sector. Establishing a multi-dimensional data collection system, creating dynamic comprehensive evaluation index models, and designing data-driven teaching feedback mechanisms can achieve precise diagnosis and scientific evaluation of teaching quality, providing strong support for teaching improvement. However, the application of data mining technology requires teachers to continuously improve their data literacy, and schools to strengthen the construction of educational information infrastructure. With the development of artificial intelligence, big data, and other technologies, data mining-based evaluation of English teaching quality will become increasingly intelligent and precise. English educators should actively embrace new

technologies and data analysis to innovate teaching reforms that bridge the gap between language proficiency and the technical complexities of the modern industry. Integrating industry-specific data into instructional practices ensures the cultivation of successors capable of engaging in precise cross-cultural communication regarding advanced fiber technologies and international manufacturing standards.

### AUTHOR CONTRIBUTIONS

Yun Cao designed, collected and analyzed the data, and drafted the manuscript. Yun Cao conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Yun Cao participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

### CONFLICTS OF INTEREST

The author declares no conflict of interest.

### FUNDING

This research was supported by,

2022's Harbin University of Commerce Teacher "Innovation" Project: Research on the Construction of Language Service System in Heilongjiang Pilot Free Trade Zone, 21YYE414

### ACKNOWLEDGEMENTS

Not applicable.

### REFERENCES

- [1] Y. Gao, "Deep learning-based strategies for evaluating and enhancing university teaching quality," *Computers and Education: Artificial Intelligence*, vol. 8, 2025, doi: 10.1016/j.caeai.2025.100362.
- [2] Z. Y. Zhang, Q. Gao, and F. Chen, "Evaluating English Language Teaching Quality in Classrooms Using OLAP and SVM Algorithms," *Mobile information systems*, vol. 2022, no. Pt.4, pp. 9327669.1-9327669, 2022, doi: 10.1155/2022/9327669.
- [3] Y. Wang, Y. Zhang, and Z. G. Dong, "Neural Network-Based Approach for Evaluating College English Teaching Methodology," *Mathematical Problems in Engineering: Theory, Methods and Applications*, vol. 2022, no. 9, pp. 1.1-1.8, 2022, doi: 10.1155/2022/2371583.
- [4] Q. L. Wang, "The Application of Big Data Fuzzy K-means Clustering in Evaluating English Teaching Ability," *Proceedings of the 2023 International Conference on Information Education and Artificial Intelligence*, pp. 318-322, 2023, doi: 10.1145/3660043.3660100.
- [5] S. S. Muttar and D. U. Al-Arnoosy, "Evaluating the Curriculum of Teaching Methods of Basic Education Colleges from the Points of View of the Teachers," *International Journal of Early Childhood Special Education*, vol. 13, no. 2, 2021, doi: 10.9756/int-jecse/v13i2.211093.
- [6] L. Cheng, L. Yu, and X. Lan, "A study on the application of data envelopment analysis in evaluating the quality of physical education teaching in universities," *International Journal of Computational Systems Engineering*, vol. 8, no. 1/2, pp. 107-119, 2024, doi: 10.1504/ijc-syse.2024.10058119.
- [7] P. Hofrichterová and D. Fait, "METHODS FOR EVALUATING TECHNICAL DRAWING AIMED AT IMPROVING THE QUALITY OF TEACHING AND INCREASING STUDENTS' LEVELS OF KNOWLEDGE," *EDULEARN Proceedings*, pp. 3584-3591, 2023, doi: 10.21125/edulearn.2023.0972.
- [8] Z. Asayesh, S. A. T. Lotfi, and S. A. H. Sarkeshikian, "Evaluating Speech Acts in English Language Teaching Software: The Case of Tell Me More," *Imam Khomeini International University*, vol. 7, no. 2, pp. 29-50, 2020, doi: 10.30479/JMRELS.2020.11660.1451.
- [9] S. Nurhidayah, A. Faridi, and Z. Shakiyya, "Evaluating the Appropriateness of English Textbook Used for Teaching English for Specific Purposes," *English Education Journal*, vol. 13, no. 1, pp. 140-151, 2023, doi: 10.15294/eej.v13i1.70649.
- [10] J. R. Chen, "A Method for Evaluating English Teaching Quality Based on Modal Data Analysis," *Learning and Analytics in Intelligent Systems*, vol. 38, pp. 81-90, 2024, doi: 10.1007/978-3-031-60777-6\_8.
- [11] C. Zhang, W. M. Cui, and X. Zhang, "A Method for Evaluating the Quality of Teaching Reform Based on Fuzzy Comprehensive Evaluation," *International Journal of Business Intelligence and Data Mining*, vol. 25, no. 3 - 4, pp. 239-252, 2024, doi: 10.1504/IJBIDM.2024.140886.
- [12] R. A. Rass, "Arab Students' Writing in English at the College Level," *Leiden, Netherlands: Brill; ro* 2024. 169-204.
- [13] B. Pokharel, "Formative Assessment in English Language Teaching," in *Giri RA, Gnawali L, editor. The Handbook of English Language Education in Nepal*. London: Routledge India, 2025.276P.
- [14] A. Sarasa-Cabezuelo and S. C. Llobet, "A Tool for Evaluating the Quality of Online Teaching," *Lecture Notes in Networks and Systems*, vol. 349, pp. 383-395, 2022, doi: 10.1007/978-3-030-90677-1\_37.
- [15] N. Li, "A Deep Learning Algorithm for Evaluating the Quality of English Teaching," *International Journal of Computational Intelligence and Applications*, vol. 22, no. 03, Art. no. 2350011, 2023, doi: 10.1142/s1469026823500116.