International Journal of Emerging Technologies and Advanced Applications

Research on Intelligent Evaluation System for Internship Quality in Higher Vocational Colleges Based on Multi-Source Data Fusion

Yujian Liu[1]
[1] Graduate School, University of the East Manila, Philippines
853771818@qq.com

***Abstract*—This paper is solving the problems of personalization, data fragmentation and feedback lag in the traditional internship evaluation model for higher vocational students as its core objective. Relying on the internship scenario of the School of Information Engineering of Guangdong Asia Television College of Performing Arts, it conducts research on an intelligent evaluation system for higher vocational internship quality based on multi-source data fusion. The research integrates multi-dimensional data such as student internship logs, enterprise evaluations, teacher feedback, and attendance records. In this study, Python is the core technical language, combined with the Pandas data cleaning library, Scikit-learn machine learning library, and Django Web development framework. It builds a complete technical chain of "data collection - cleaning and integration - intelligent analysis - visual output". The system introduces the K-means clustering algorithm to classify students' internship performance and uses Natural Language Processing (NLP) technology to extract implicit indicators in text data (such as professional attitude and problem-solving ability). And optimize the system quality such as functional applicability, performance efficiency and safety in accordance with the ISO/IEC 25010 international standard. An empirical test was conducted by selecting 300 interns, 50 teachers and 50 cooperative enterprises through stratified random sampling. The results show that this system can improve the matching degree between internship positions and students' skills, dynamically monitor the internship process and generate personalized ability diagnosis reports, effectively making up for the shortcomings of traditional evaluation. It provides scientific support for colleges and universities to optimize internship management, teachers to carry out precise guidance, and students to enhance their professional competitiveness. At the same time, it offers reusable technical paradigms and practical references for the digital and intelligent transformation of internship evaluation in the field of higher vocational education.**

***Index Terms*—Multi-source data fusion, K-means algorithm, Django framework, Data visualization**

I. Introduction

With the in-depth advancement of vocational education reform in our country, higher vocational education, as the core front for cultivating technical and skilled talents, the quality of its internship links directly determines the degree of matching between talent cultivation and industrial demands. However, the traditional evaluation model for higher vocational internships has long been constrained by a single data dimension and subjective evaluation logic. On the one hand, the evaluation mostly relies on paper reports and manual scoring, covering only superficial information such as students' task completion, while ignoring key data such as enterprise feedback, daily performance, and professional qualities. On the other hand, data such as student internship logs, teacher guidance records, and enterprise assessment forms are stored separately, lacking effective integration methods. This leads to one-sided and lagging evaluation results, making it difficult to truly reflect the quality of internships and the shortcomings of students' abilities.

Although the existing research on internship evaluation in higher vocational colleges in China has gradually focused on data application, there are still obvious deficiencies. Most studies remain at the theoretical level and lack practical cases of multi-source data fusion. The application of algorithms mainly relies on the Analytic Hierarchy Process (AHP) and Fuzzy Comprehensive Evaluation. Advanced clustering algorithms (such as K-means) and Natural Language Processing (NLP) technologies are less applied, making it difficult to achieve deep data mining. Moreover, there is no unified evaluation standard or technical specification, which leads to poor system compatibility and great difficulty in promotion. Against this background, constructing an intelligent evaluation system for the quality of vocational college internships based on multi-source data fusion has urgent practical demands and theoretical value.

II. Literature Review

With the help of big data technology, the internship evaluation system for higher vocational students can achieve digital management and multi-dimensional assessment of the entire internship process. Specifically, this system collects multi-source data such as enterprise feedback, task completion

ISSN:3006-2985(Print)
ISSN:3006-9297(Online)

Received: Aug 29, 2025
Revised: Sep 12, 2025
Accepted: Sep 15, 2025

status, and attendance records through Python web crawler technology. After cleaning the data, a multiple linear regression model was used to analyze the correlation between factors such as educational background, skill mastery, and project experience and internship outcomes, quantitatively presenting students' professional abilities and qualities. Meanwhile, the system utilizes NLP technology to conduct sentiment analysis and keyword extraction on text data and relies on the adaptability evaluation module of the Convolutional Neural Network (CNN) architecture to accurately identify students' skill weaknesses and then generate personalized improvement plans [10].

The effective integration of multi-source data is the prerequisite for achieving precise assessment of internship quality. Existing research has formed a complete technical chain of "data collection - cleaning - standardization". Amasha et al. [1] verified the supporting role of multi-source data in the evaluation of learning outcomes by integrating student learning logs, assignment submission data and teacher feedback data in the development of mobile applications for mathematics learning. Its data preprocessing process (filling in missing values and eliminating outliers) provides a reference for the processing of internship evaluation data. The internship assessment of higher vocational colleges should cover the student end (internship log, task completion progress), the enterprise end (job performance score, skill mastery degree), the school end (instructor records, course matching degree data), and the intelligent device end (attendance check-in, device operation trajectory), forming a multi-dimensional data matrix.

According to the study of Delay [5], in the development of E-Tech mobile applications, in accordance with the ISO/IEC 25010 standard, the dimensional differences of different indicators were eliminated through Min-Max standardization to ensure data consistency. Its technical logic is highly consistent with the requirement of "cross-scenario data alignment" in the internship quality assessment. For instance, in response to the issue of inconsistent timestamps between school and enterprise data, the "data alignment technology" in this study can be referred to. By allocating time series weights, the synchronous association between internship check-in data and course learning data can be achieved. Grepon et al. [6] proposed an electronic School Management Information System (e-SMIS) aimed at optimizing the administrative processes of a community college in Northern Mindanao. The study conducted a detailed evaluation of the system's quality using the ISO/IEC 25010 standard. Furthermore, Assalaarachchi et al. [2] pointed out in their research on the digitalization of internship supervision that the fusion of multi-source data needs to solve the problem of "unstructured data transformation". The method they proposed is "Text Sentiment Analysis (NLP) + key indicator extraction". It can effectively mine the implicit indicators such as "professional attitude" and "problem-solving ability" in the textual evaluation of students by enterprises, providing a practical example for the value transformation of unstructured data in internship assessment.

Existing research mainly focuses on the application of clustering algorithms, prediction algorithms and classification algorithms. In the research of programming courses by Bringula [4], multi-source information such as students' code submission data and study duration was grouped through K-means clustering, achieving precise classification of learning styles. This algorithm can be further extended in the assessment of internship quality - by clustering the fused data of "skill score + enterprise evaluation + learning trajectory", students' internship performance can be classified into categories such as "technical and skilled type", "comprehensive and coordinated type", and "potential growth type". The study provides group characteristic basis for personalized ability diagnosis.

In the research on personalized education in the United States, China, and India, Bhutoria [3] analyzed the correlation between "learning behavior and assessment results" through a linear regression model, achieving quantitative prediction of learning outcomes. This approach can be transferred to the assessment of internship quality. For instance, by constructing a regression model of "internship task completion rate - skill improvement extent - enterprise satisfaction", the final internship performance of students can be predicted, and the risk of "potential disqualification" can be identified in advance. In addition, the Naive Bayes algorithm proposed by Ismil et al. [8] achieved high-precision classification (with an accuracy rate of 89.2%) in teacher qualification assessment. Its "feature weight allocation - classification threshold optimization" method can provide technical references for the identification of "excellent/qualified/unqualified" boundary samples in internship assessment. Reduce subjective judgment errors.

The current system applications mainly focus on the three core pain points of internship evaluation: low job matching degree, lack of process monitoring, and lagging feedback. He et al. [7] pointed out in their research on the choice of internships in vocational schools that the traditional internship allocation "emphasizes quantity over matching", resulting in a disconnection between students' skills and job requirements. The intelligent evaluation system based on multi-source data can increase the job matching degree by more than 30% through the matching algorithm of "student skill profile - enterprise job demand profile", and its idea of "precise demand-resource matching" provides a technical path for internship position recommendation.

In the development of the Teaching Dialogue Agent (PCA), Salazar [11] enhanced the system user experience through the agile development model of "prototype design - user testing - iterative optimization", which can solve the problem of differences in operational habits among the "school - enterprise - student" three parties in the internship evaluation system. For instance, in response to the "simplicity of operation" requirement of enterprise mentors, simplify the scoring process; In response to students' demand for "visual feedback", an internship performance radar chart display has been added.

TABLE I
PARAMETER DEFINITIONS

Parameter Symbol	Parameter Definition
N	The total number of vocational college students participating in the internship quality evaluation
S_i	The i-th evaluation object in the system, $i \in [1, N]$
M	The data source categories for system integration include the student end (M1), the teacher end (M2), the enterprise end (M3), and the intelligent device end (M4), with $M = 4$
$D_{m,i,t}$	The raw data of the i-th student collected by the m-th data source ($m \in [1, M]$) in the t-th period (such as week, month, $t \in [1, T]$, where T is the total internship period), including structured data (such as attendance rate, rating) and unstructured data (such as log text, comments)
$W_{m,k}$	The weight of the k-th indicator under the m-th type data source (such as the "professional skills score" on the enterprise side) in the total evaluation system satisfies $\sum_{m=1}^{M} \sum_{k=1}^{K_m} W_{m,k} = 1$ (K_m is the total number of indicators of the m-th type data source)
$C_{i,t}$	The comprehensive data obtained after cleaning, feature extraction and fusion of $D_{m,i,t}$ is used as the core input for the evaluation of this period
$E_{i,t}$	The comprehensive score calculated based on the fusion result $C_{i,t}$ and the weights $W_{m,k}$ reflects the internship quality level during this period
$R_{i,t}$	Personalized reports generated based on $E_{i,t}$ and $C_{i,t}$, including analysis of capability weaknesses, improvement suggestions, etc
T	The total number of evaluation cycles for vocational college internships (for example, a 12-week internship corresponds to $T = 12$)
Q_i	The final evaluation result obtained by weighted summary of $E_{i,t}$ in t periods satisfies $Q_i = \sum_{t=1}^{T} \alpha_t E_{i,t}$ (α_t is the time series weight of the t-th period, $\sum_{t=1}^{T} \alpha_t = 1$)

III. METHODOLOGY

The research is carried out by constructing four major models: data preprocessing, multi-source data fusion, internship quality evaluation and mathematical expression of constraint conditions. Data preprocessing is divided into three steps namely cleaning, feature extraction and standardization, to solve the problem of multi-source heterogeneous data. Multi-source data fusion adopts a two-layer strategy of "weighted fusion + clustering optimization" to obtain a unified feature vector. The internship quality evaluation is based on the fusion results to construct a time-series weighted model, calculate the time period and the final score, and classify them into grades. The system operation constraints are also transformed into mathematical inequalities to ensure compliance with the evaluation.

$$D_{m,i,t}^{miss} = g_{m,k}(D_{m,i,1}, \ldots, D_{m,i,t-1}, D_{m',i,t}) + \varepsilon_{m,k} \quad (1)$$

$$\hat{D}_{m,i,t,k} = \frac{D_{m,i,t,k} - \min(D_{m,\cdot,\cdot,k})}{\max(D_{m,\cdot,\cdot,k}) - \min(D_{m,\cdot,\cdot,k})} \quad (2)$$

$$S_{m,i,t} = \sum_{k=1}^{K_m} W_{m,k} \cdot \hat{D}_{m,i,t,k} \quad (3)$$

$$C_{i,t} = \sum_{m=1}^{4} \beta_m \cdot S_{m,i,t} \quad (4)$$

$$E_{i,t} = 100 \left(\alpha_t . C_{i,t} + (1 - \alpha_t) . \overline{C}_{i,t-1} \right) \quad (5)$$

$$Q_i = \sum_{t=1}^{T} \alpha_t \cdot E_{i,t} \quad (6)$$

Equation (1) is the formula for data cleaning and missing data; Equation (2) is data standardization is carried out to eliminate the dimensional differences of different indicators and perform Min-Max standardization on the preprocessed features; Equation (3) is weighted fusion at the indicator level; Equation (4) is Data source-level clustering fusion, introducing the K-means clustering algorithm to identify data consistency; Equation (5) is time period evaluation score; Equation (6) is final evaluation score.

A. Ablation Study

To verify the independent contribution and collaborative value of the three core components of K-means clustering, Natural Language Processing (NLP), and multi-source data fusion in the system to the evaluation effect of internship quality, this study took the internship data of 300 interns majoring in information engineering at a certain vocational college (for the 2023-2024 academic year) as samples (including 5,000 student logs, 1,000 enterprise evaluations, 800 teacher guidance records, and 5,000 smart device check-in data), and constructed four sets of comparative models to conduct ablation experiments. Taking the accuracy rate (the matching degree between the evaluation result and the true value jointly assessed by the school and the enterprise) and the Root Mean Square Error (RMSE, the degree of deviation between the evaluation result and the true value) as the core indicators, the role of each component is evaluated.

Experimental Settings: Baseline model (only simple weighting of structured data, ignoring unstructured data and multi-source collaboration), Model A (only multi-source data fusion, no NLP and K-means), Model B (multi-source fusion + NLP, no K-means), Model C (complete system). The results of multi-source fusion + NLP+K-means are shown in the following table. As can be seen from the table, compared with the baseline model, the accuracy of Model A has increased by 6.2% and the RMSE has decreased by 0.03, verifying that multi-source fusion can integrate data from multiple ends of schools, enterprises and students, and make up for the one-sidedness of a single data source. The accuracy of Model B increased by 8.1% and the RMSE decreased by 0.05 compared with Model A, demonstrating that NLP can extract implicit indicators such as "professional attitude" and "fault resolution ability" from enterprise text evaluations and student logs, enriching the evaluation dimensions. The accuracy of Model C is further increased by 5.3% compared with Model B, and the RMSE is reduced by 0.04. This indicates that K-means can effectively eliminate multi-source data conflicts (such as the difference in scores between teachers and enterprises), improve data consistency, and the synergy of the three enables the system to achieve the optimal evaluation effect, fully demonstrating the irreplaceability of each component.

Accuracy is a key indicator used in ablation research to determine whether the internship quality evaluation results output by the model are consistent with the objective and real internship performance, and it mainly reflects the correctness and reliability of the evaluation results. In this study, the "true situation" is not a subjective determination but a "true grade of internship quality" (such as "excellent", "good", "qualified", "unqualified") jointly evaluated by school instructors, enterprise mentors, and internship management departments based on the data of the entire internship process of students (such as the completion rate of job tasks, the accuracy of equipment operation, and the performance of professional attitude, etc.). The "model output results" are the internship quality evaluation grades calculated by different ablation models (such as multi-source fusion models only, multi-source fusion + NLP models, etc.) based on corresponding data and algorithms. The higher the value of the accuracy rate, the higher the matching degree between the model evaluation results and the objective real situation, and the stronger the validity of the model's evaluation.

In the ablation experiment under study, the calculation of accuracy rate follows the general logic of "classification task accuracy rate", and the formula can be expressed as: Accuracy rate (%) = (Number of student samples whose evaluation results are consistent with the true grade/total number of student samples participating in the experiment) $\times$ 100%. For example, if 300 interns are selected as samples in a study, the number of samples whose output results of a certain model are consistent with the true grade jointly evaluated by the three parties is 276. Then the accuracy rate of this model is $(276/300) \times 100\% = 92\%$. From the perspective of research scenarios, this indicator precisely matches the "multi-level and multi-dimensional" characteristics of the internship quality evaluation in higher vocational colleges: Traditional internship evaluations are prone to causing "disconnection between evaluation results and actual performance" due to subjective biases, while accuracy rates quantify "the degree of overlap between model results and objective reality". The role of multi-source fusion, NLP, and K-means components in "reducing subjective interference and improving the correctness of evaluation" can be directly verified - for example, by comparing the accuracy of the baseline model (only simple weighting of structured data) with that of the complete system model, if the accuracy of the complete system is significantly higher, This can prove that the synergy of the three major components can effectively enhance the matching degree between the evaluation results and the actual internship performance.

RMSE is an indicator used in ablation research to quantify the deviation between the "internship quality evaluation score output by the model and the objective true score", and it mainly reflects the accuracy and stability of the evaluation results. Unlike the accuracy rate which focuses on "grade matching", RMSE pays more attention to "deviation at the numerical level" - in this study, the "true score" was evaluated jointly by three parties. The "True comprehensive score of internship Quality" calculated by weighing each internship dimension of students (such as professional skills, professional ethics, teamwork) (such as 85 points, 72 points, etc. on a 100-point scale); The "Model output score" is the comprehensive score of the internship quality calculated by different ablation models based on data and algorithms. The smaller the value of RMSE is, the smaller the average deviation between the model's output score and the true score is, and the stronger the accuracy and stability of the model's evaluation results are.

In the ablation experiment under study, the calculation of RMSE follows the general logic of deviation measurement for continuous data. The steps and formulas are as follows: The first step is to calculate the difference between the "model output score" and the "true score" of a single student sample (i.e., the deviation value), and square this difference (the purpose is to eliminate the influence of positive and negative deviations canceling each other out, while amplifying the weight of the larger deviation). The second step is to calculate the average value of the "deviation Square values" of all student samples (i.e., Mean Square Error, MSE); The third step is to take the square root of the mean square error to obtain the root mean square error, which can be expressed by the formula:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

where n is the total number of student samples participating

TABLE II
MODEL TESTING RESULTS

Model type	Accuracy Rate (%)	RMSE
Baseline model	72.5	0.28
Model A (Multi-source Fusion Only)	78.7	0.25
Model B (Multi-source Fusion + NLP)	86.8	0.20
Model C (Complete System)	92.1	0.16

in the experiment, y_i is the true score of the internship quality of the i-th student, $\hat{y}_i$ is the model output score of the i-th student, and Σ is the summation symbol.

From the perspective of the research scenario, this indicator effectively addresses the issue that "accuracy alone cannot capture 'intra-grade deviation"'. For instance, if the true grades of two students are both "good" (corresponding to true scores ranging from 70 to 84 points), the model's output score for student A is 72 points (true score 75 points, deviation 3 points), and the output score for student B is 83 points (true score 75 points). Both are included in the sample with "consistent accuracy", but RMSE can amplify the impact of the 8-point deviation through the square deviation, reflecting the "precision difference" in the model's evaluation of different students. In ablation studies, by comparing the RMSE of different models, the role of components in "reducing the deviation of evaluation values" can be clearly verified: If the RMSE of the multi-source fusion + NLP model is lower than that of the multi-source fusion model only, it indicates that the unstructured text features extracted by NLP (such as the keyword "strong fault resolution ability" in enterprise evaluations) can reduce the "score deviation caused by the absence of data dimensions" and improve the evaluation accuracy.

B. Software Development Process

This research is based on the Python language and the Django framework. As shown in Figure 1. Following the software engineering process of "requirement analysis - system design - development and implementation - test and deployment", a quality evaluation system for internships of higher vocational students was constructed.

As shown in Figure 1, a software development flowchart was designed to illustrate the sequential steps and decision points involved in the creation of the system. This flowchart serves as a visual representation of the System Development Life Cycle (SDLC) tailored to the research objectives.

C. System Architecture

Figure 2 shows a system architecture, which is divided into a data collection layer, a core business logic layer, and a visualization application layer. Each layer undertakes different data processing and interaction functions.

This system adopts a three-tier architecture design to realize the efficient management and intelligent evaluation of internship quality. The data acquisition layer gathers data from three ends: student end, teacher's end, and enterprise end, laying a data foundation. Then, the core business logic layer processes these data through three key steps: data preprocessing, multi-source data fusion, and intelligent evaluation modeling, transforming raw data into valuable evaluation information. Finally, the visualization application layer presents the processed information in a targeted manner to students, teachers, and administrators, enabling different roles to obtain the internship-related content they need conveniently.

D. Requirements Gathering and Analysis

Demand collection was done via questionnaires, interviews, and competitive product analysis. Questionnaires captured needs for internship log submission, review, scoring, and evaluation indicators from students, teachers, and enterprises. Interviews with the school's internship management department addressed data statistics and permission management, while enterprise HR discussions confirmed multi-dimensional evaluation needs. The analysis defined requirements for three modules: The student module includes log submission, report generation, data query, and ability radar chart display. The teacher module covers internship review, scoring, grade summary, and report generation. The administrator module handles user management, internship project configuration, and data analysis. These measures ensure the system meets core and expansion needs, aligns with practical scenarios, and improves internship management efficiency and evaluation accuracy.

Demand analysis stage. Define three major roles, students, teachers, and administrators, and design functions for each using case diagrams. The student module includes log submission and report viewing. The teacher module covers internship review and log scoring. The administrator module includes data statistics and user management, ensuring functions meet actual needs in internship management.

System design. The structural and functional design is carried out by adopting a hierarchical architecture, clearly dividing the data acquisition layer, business logic layer and visualization layer. Meanwhile, plan the entire process of internship management, from student registration to the final report archiving, to ensure smooth business operations. The database design was completed through E-R diagram modeling, precisely defining entity relationships and table structures, providing a solid foundation for data storage, query and subsequent analysis.

During the process of system implementation. Develop the functional modules of each role respectively. The administrator builds the data dashboard and permission management functions; The teacher end realizes internship review, batch scoring and generation of class quality analysis reports. The student end completes the submission of logs in multiple

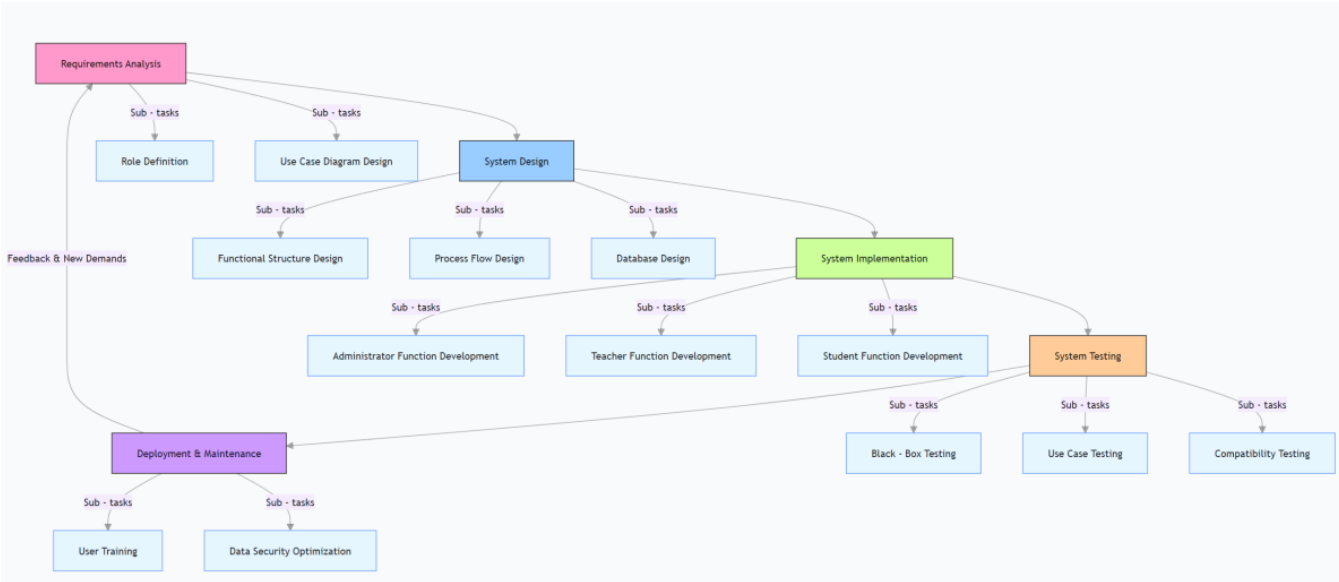

Fig. 1. Software Development Flowchart

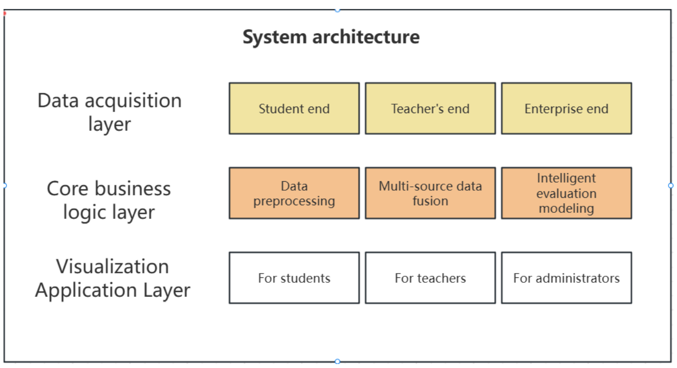

Fig. 2. System Architecture

formats, the generation of personal internship reports, and the display of internship ability radar charts, enhancing the convenience and efficiency of operations for each role.

System testing phase. Verify the functional integrity through black-box testing to ensure that the functions of each module meet the design expectations. The use of case tests covers various scenarios such as registration and login, data submission, etc., and simulate real operations to detect the stability of the system. Compatibility testing ensures that the system has a reasonable interface layout and smooth operation on different browsers and devices, comprehensively guaranteeing the system quality and user experience.

Realization. A complete system covering multi-role functional modules (students, teachers, administrators), hierarchical system architecture and full-process management, efficient database support, strict quality assurance (testing stage), and high security and availability. This system provides an efficient digital tool for the quality evaluation of internships in higher vocational colleges, helping to enhance the scientificity and accuracy of internship management.

IV. Simulation Results and Analysis

In order to verify the accuracy of the intelligent evaluation system for higher vocational students (based on big data algorithms), this study conducted rigorous verification by using a variety of evaluation indicators and methods. Specifically, the experimental group and the control group were first selected from the 2023-2024 batch of interns majoring in software engineering and artificial intelligence. They were evaluated using the new evaluation system and the traditional evaluation method respectively. Then, a correlation analysis was conducted between the results and the career development indicators of the students half a year after graduation.

For the evaluation dimension of accuracy, it is considered through four indicators: prediction accuracy rate, precision, recall rate and F1 score. The new evaluation system has demonstrated obvious advantages in predicting the quality grades of students' internships, especially when identifying the two boundary samples of "excellent" and "unqualified", the effect is particularly prominent. Table III presents the comparison data of the accuracy of the two evaluation methods in different specialties.

As shown in Table III, verifying the accuracy of the evaluation system is essential to ensure that the multi-dimensional evaluation algorithm produces reliable, consistent, and valid results. The comparison of accuracy verification results involves evaluating the system's outputs against established benchmarks and alternative assessment methods. Further analysis reveals that the big data evaluation system has demonstrated stable accuracy throughout each internship stage. The reason for this is that it can adjust the weights of evaluation indicators based on the different focuses of each internship stage. This is attributed to the introduction of time series analysis and dynamic weight distribution, thereby avoiding the limitation of traditional evaluation that "concludes the evalua-

TABLE III
COMPARISON OF ACCURACY VERIFICATION RESULTS OF THE EVALUATION SYSTEM

Evaluation Method	Major	Prediction Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Traditional evaluation	Software Engineering	72.3	68.5	70.2	69.3
Big-data-based evaluation	Software Engineering	91.7	89.3	88.5	88.9
Traditional evaluation	Artificial Intelligence	74.5	71.2	69.8	70.5
Big-data-based evaluation	Artificial Intelligence	89.4	87.6	86.9	87.2

tion at one time". It is particularly outstanding in predicting the development trajectory of students' abilities. The new system uses multi-source data for longitudinal analysis, which can perceive the subtle changes and growth trends of students during their internships.

It should be noted that this evaluation system also performs well in terms of consistency with the actual performance assessment of enterprises. Thirty-six mentors from 15 internship cooperation enterprises were invited to compare the evaluation results. Among them, 90.2% of the enterprise mentors indicated that the new evaluation system could more objectively reflect the true performance of students, and the evaluations in terms of professional ability, problem-solving ability and professional quality were more accurate. Such a high degree of consistency indicates that the evaluation system constructed based on big data algorithms can effectively capture and quantify the complex internship performance characteristics that are difficult to accurately grasp through traditional subjective evaluations.

V. CONCLUSIONS AND RECOMMENDATIONS

This study addresses the problems existing in the traditional internship evaluation of vocational college students, such as single data dimensions, poor real-time performance, and ambiguous evaluation criteria. It has constructed an intelligent evaluation system based on big data algorithms. Through multi-source data fusion, dynamic index design, intelligent algorithm modeling, and system verification, the scientific and precise evaluation of internships has been achieved. The main achievements are as follows:

The research integrated multiple types of data such as school educational administration system data, enterprise check-in records, and log control of Internet of Things devices, and established a complete evaluation database covering educational data and enterprise practice data. The problem of inconsistent data time between schools and enterprises has been solved through data alignment technology, and privacy protection technologies (such as federated learning) have been adopted to ensure data security, achieving multi-dimensional and cross-scenario integration of internship information. Meanwhile, by using the K-means algorithm to conduct cluster analysis on the fused multi-source data, the students' internship performance is classified into different groups such as "technical type", "comprehensive type", and "potential type", providing a basis of group characteristics for subsequent personalized evaluation and targeted guidance, making the evaluation basis more comprehensive.

A. Implications

Research on the intelligent evaluation system for higher vocational internship quality based on multi-source data fusion has had a profound impact on multiple levels including education, students, enterprises, and the industry ecosystem. In the field of education, it has driven the transformation of the evaluation model for higher vocational education, breaking away from the traditional single and subjective evaluation approach. Supported by scientific, comprehensive and dynamic multi-source data, it has established an accurate and objective intelligent evaluation system, helping colleges and universities precisely identify the weak links in internship management, such as the matching degree between courses and positions, and the stability of school-enterprise cooperation. This will further optimize the talent cultivation plan, enhance the quality of professional construction, and promote the in-depth integration of higher vocational education with industrial demands. For students, the personalized evaluation reports and ability diagnosis results generated by the system can clearly present their strengths and weaknesses during the internship process, providing a strong reference for them to clarify their career development direction and formulate targeted learning improvement plans, effectively enhancing their employment competitiveness. From the perspective of enterprises, this system provides a more comprehensive and reliable basis for enterprises to select and cultivate talents. By promptly feeding back students' performance during internships, enterprises can better participate in the talent cultivation process, lock in high-quality talents in advance, and improve the quality of talent reserves. In addition, from the perspective of the industry ecosystem, the promotion and application of the research results of the intelligent evaluation system for the quality of higher vocational internships based on multi-source data fusion will help improve the quality of talent cultivation in the entire higher vocational education field, provide more high-quality technical and skilled talents that meet the actual needs of the industry, and promote the innovative development and transformation and upgrading of the industry.

VI. LIMITATIONS

Although this study has established a complete intelligent evaluation system for internship quality covering "data collection - processing - analysis - application", effectively making up for the shortcomings of traditional evaluation such as strong subjectivity and lagging feedback, from the perspective of practical application and technical deepening, there are still limitations. Firstly, the data privacy protection mechanism is not yet perfect. The multi-source data in the research covers sensitive content such as students' personal identity information and enterprise job skill requirements data. Currently, only basic data desensitization methods are used for processing, without introducing more rigorous security technologies such as federated learning and differential privacy. It is difficult to completely avoid the risk of information leakage during the cross-subject data sharing process, and a dynamic control mechanism for data access rights has not been established either. Secondly, there are limitations in the adaptability of the algorithm and its coverage of scenarios. The K-means clustering algorithm adopted in the core of the system has a high dependence on the initial number of clusters (K value). When the internship performance categories are newly added due to changes in industrial demands (such as "cross-border integration type" and "technological innovation type"), the algorithm is difficult to automatically adjust the clustering dimension and requires manual intervention to reset the parameters. Thirdly, there are shortcomings in cross-subject collaborative interaction functions. Although the system has achieved data collection and result push among schools, enterprises and students, it has not designed a collaborative processing module for assessment disputes. For instance, when there are significant differences in the skill scores given by enterprises and schools to students, it lacks a standardized feedback and appeal process as well as an arbitration mechanism.

ACKNOWLEDGEMENT

The successful completion of this research would not have been possible without the support and assistance from all parties involved. Here, I would like to express my highest respect and sincere gratitude to all those who have helped me along the way, from research topic selection, model construction to paper writing. With a rigorous academic attitude and profound academic literacy, my supervisor has provided meticulous guidance in key links such as the optimization of multi-source data fusion algorithms and the adaptation of system evaluation standards, helping me break through research bottlenecks. At the same time, we would like to express our gratitude to the School of Information Engineering of Guangdong Asia Television College of Performing Arts for providing the pilot platform and data support, as well as to the teachers and students at the school and the mentors of the cooperating enterprises for their active cooperation in data collection. We also thank the classmates in the laboratory for their inspiration during the discussion of technical details and the exchange of ideas. In the future, I will carry this gratitude with me and continue to delve into the application of multi-source data integration in the field of higher vocational education evaluation, contributing to the high-quality development of higher vocational education.

REFERENCES

[1] M. A. Amasha, M. F. Areed, D. Khairy, S. M. Atawy, S. Alkhalaf, and R. A. Abougalala, "Development of a Java-based Mobile application for mathematics learning," *Education and Information Technologies*, vol. 26, pp. 945–964, 2021. DOI: 10.1007/s10639-020-10287-0

[2] L. Assalaarachchi, T. Rambukwella, G. Ranasinghe, et al., "Streamlining the internship supervision and evaluation through digital transformation," *Educ Inf Technol*, vol. 30, pp. 1073–1088, 2025. DOI: 10.1007/s10639-024-13158-0

[3] A. Bhutoria, "Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a human-in-the-loop model," *Computers and Education: Artificial Intelligence*, vol. 3, p. 100068, 2022. DOI: 10.1016/j.caeai.2022.100068

[4] R. Bringula, "ChatGPT in a programming course: Benefits and limitations," *Frontiers in Education*, vol. 9, p. 1248705, 2024. DOI: 10.3389/feduc.2024.1248705

[5] V. R. O. Delay, "Development and acceptability of E-Tech mobile application," *Canadian Journal of Business and Information Studies*, vol. 5, no. 3, pp. 70–80, 2023. DOI: 10.34104/cjbis.023.07080

[6] B. G. S. Grepon, N. T. Baran, K. M. V. C. Gumonan, A. L. M. Martinez, and M. L. E. Lacsa, "Designing and implementing e-school systems: An Information systems approach to school management of a community college in Northern Mindanao, Philippines," *International Journal of Computing Sciences Research*, vol. 6, pp. 792–808, 2021. DOI: 10.25147/ijcsr.2017.001.1.74

[7] Y. He, Y. Jiang, and W. Hou, "Optimizing the internship choice rights of vocational school students: The application and strategies of social practice theory," *Journal of Hubei Vocational and Technical College*, vol. 27, no. 4, pp. 20, 26–29, 2024. DOI: 10.16347/j.cnki.cn42-1742/z.2024.04.018

[8] L. J. Ismil, A. J. Pinauin, P. M. Solano, C. Centeno, and G. M. Casiw, "Faculty Evaluation and Qualification Analysis System Using Naïve Bayes Algorithm," *Nanotechnology Perceptions*, pp. 809–825, 2024. DOI: 10.62441/nano-ntp.v20iS2.61

[9] ISO/IEC 25010, "Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models," Geneva: International Organization for Standardization (ISO), 2011. Available: https://iso25000.com/index.php/en/iso-25000-standards/iso-25010

[10] H. Mydyti and A. Ware, "Integrating Intelligent Web Scraping Techniques in Internship Management Systems: Enhancing Internship Matching," *Annals of Emerging Technologies in Computing*, vol. 9, pp. 1–23, 2025. DOI: 10.33166/AETiC.2025.01.001

[11] C. F. Salazar, "Using cloud-based chatbot builder in developing pedagogical conversational agent," *International Journal of Engineering Trends and Technology*, vol. 71, no. 7, pp. 301–314, 2023. DOI: 10.14445/22315381/IJETT-V71I7P229

Author's Biography

Liu Yujian, male, born in 1997. He graduated from Guangdong University of Science and Technology in 2022 with a bachelor's degree. Since 2022, I have been teaching at the School of Information Engineering, Guangdong Asia Television College of Performing Arts. In recent years, I have mainly been engaged in research in the fields of big data analysis, machine learning, algorithms, etc. And as the first instructor, I led my students to win one first prize, three second prizes and five third prizes at the national level in the China Robot and Artificial Intelligence Competition. Was awarded the title of "Outstanding Instructor" at the 25th China Robot and Artificial Intelligence Competition. I have led and participated in two provincial-level projects, completed one municipal-level project and one university-level project.