Construction and Empirical Study of a Personalized Teaching Model for Vocational College Music Performance Major Based on Intelligent Recommendation Algorithm

Yixin Cao, Na Li, Chenlu Liu, Ping Wang, Nina Gao, Jiangbo Guo, Wen Song

School of Culture, Tourism, Health Care and Education, Dongying Vocational College, Dongying 257000, Shandong, China

Corresponding author: Wen Song (e-mail: wensong2026123@163.com)

ABSTRACT In networked audio-visual teaching environments, the processing and transmission of vocal and instrumental performance signals require reliable feature extraction, low-latency communication, and adaptive decision mechanisms that are consistent with engineering approaches used in intelligent sensing and wireless communication systems. This study addresses the substantial differences in skill level, learning style, and development direction among vocationalcollege music performance students by constructing a personalized teaching recommendation system that integrates collaborative filtering with content-based filtering. Student performance behavior, vocal practice records, and instrumental audio features are collected and transformed into multidimensional learner profiles, enabling the system to generate differentiated teaching-resource allocation schemes and dynamic practice paths. A quasi-experimental design was implemented with two parallel classes in a vocational-college music performance program, and teaching effectiveness was evaluated through skill achievement, learning autonomy, and resource-utilization efficiency. The results show that the proposed model significantly improves students' vocal and performance skill acquisition, enhances learning motivation, and increases the matching accuracy of teaching resources. The study provides an operational paradigm for intelligent art-teaching transformation and offers a data-driven reference for adaptive signal-based learning systems.

INDEX TERMS intelligent recommendation algorithm,music performance education,acoustic feature extraction,wireless communication,performance skill assessment

I. INTRODUCTION

OWING to its emphasis on practicality and individual artistic expression, the music performance major in vocational colleges presents a high degree of heterogeneity among students in terms of entrance foundation, skill structure and career development intention. The traditional class teaching system advances teaching with a unified progress and unified content, which is difficult to take into account the significant individual differences among students, leading to the long-term existence of a polarization phenomenon where some students are "not challenged enough" while others "cannot keep up". The development of educational big data and machine learning technologies has enabled intelligent recommendation systems to provide technical support for solving this dilemma, and the dynamic matching of teaching resources and learning tasks can be realized through mechanisms such as collaborative filtering and content-based filtering. However, the existing algorithm models still have limitations in adapting to unstructured data such as performance audio and vocal scores in the field of music performance [1], and whether personalized recommendation can be effectively transformed into skill improvement effects still needs to be verified through rigorous empirical research.

II. THEORETICAL FOUNDATION AND ADAPTIVE TRANSFORMATION OF RECOMMENDATION ALGORITHM FOR MUSIC PERFORMANCE MAJOR

A. WEIGHT CORRECTION MECHANISM OF COLLABORATIVE FILTERING ALGORITHM IN MUSIC SKILL SEQUENCE LEARNING

The core logic of the collaborative filtering algorithm is to infer individual preferences by using the historical behavior data of user groups, but when it is transplanted to the music skill learning scenario, the original algorithm's practice of assigning similar weights to all learning records has obvious defects. The acquisition of music skills has a strict sequence dependence — there is a pre-constraint relationship among breath control, intonation stability and phrase expression, resulting in a significant attenuation of the predictive power of early practice records on the current skill state [2]. Therefore, it is necessary to introduce a dual weight correction of time attenuation and skill level in the similarity calculation link of the user-resource interaction matrix. Let the interaction score of learner u for resource i at time t be $\tilde{r}_{ui}$, the corrected effective score is defined as:

$$\tilde{r}_{ui} = r_{ui}^{t} \cdot e^{-\lambda(T-t)} \cdot \omega_l \tag{1}$$

where λ is the time attenuation coefficient, T is the current time, and ω_l is the level weight corresponding to the skill level l to which the resource belongs. The user similarity is recalculated based on the corrected scores, so that peer neighbors with similar skill stages gain higher influence in recommendation. In the actual deployment of the vocational college music performance scenario, the parameters were calibrated through a preliminary sensitivity analysis on a historical dataset of 500 practice records. The time attenuation coefficient λ was optimized to 0.08 to maximize the Mean Reciprocal Rank (MRR) within a standard 16-week semester cycle. Similarly, ω_l values were empirically derived as 0.6, 1.0, and 1.4 for the basic, intermediate, and comprehensive levels respectively. Comparative testing indicated that this specific weight distribution, compared to uniform weighting, improved the Precision@5 of recommendations by 12.4%, thus making the recommendation results incline to the practice resources matching the learner's current skill stage and effectively avoiding the problem of cross-stage mismatched recommendation [3].

B. MODELING OF CONTENT-BASED FILTERING ALGORITHM BASED ON VOCAL AND INSTRUMENTAL MUSIC CONTENT FEATURES

The content-based filtering algorithm recommends similar content to learners according to the inherent attributes of resources, and the key difficulty in its application in the field of music performance is how to transform unstructured audio data into computable feature vectors. There are significant differences between vocal and instrumental music in the core characterization dimensions: vocal music resources need to focus on depicting vocal tract physiological features such as vocal range span, passaggio position, articulation clarity and breath stability, while instrumental music resources rely more on timbre spectrum envelope, technique key touch force curve and rhythm density distribution [4]. For vocal music resources, 13-dimensional timbre features are extracted by using Mel Frequency Cepstral Coefficients (MFCC), and the pitch trend sequence is obtained by combining the pitch detection algorithm, together forming a 48-dimensional acoustic feature vector; for instrumental music resources, the spectral centroid and roll-off frequency of the Short-Time Fourier Transform are additionally introduced to expand the feature dimension to 64. The feature vectors of the two types of resources are normalized by L2, and the content affinity between resources is measured by cosine similarity, thus constructing a classified resource-resource similarity matrix [5]. This separate modeling strategy avoids the cross-type semantic distortion caused by mixing vocal and instrumental music into a unified feature space, and keeps the content recommendation results accurately aligned in the professional direction.

C. FUSION LOGIC OF HYBRID RECOMMENDATION STRATEGY AND DOMAIN-SPECIFIC SOLUTION TO COLD START PROBLEM

Collaborative filtering relies on the accumulation of interaction data, while content-based filtering relies on the characterization of resource features, and the two form a complementary relationship in terms of accuracy and coverage. The hybrid recommendation strategy obtains comprehensive advantages by weighted fusion of the two sets of scores. Let the predicted score of collaborative filtering for learner u on resource i be $\hat{r}_{ui}^{CF}$, and the predicted score of content-based filtering be $\hat{r}_{ui}^{CB}$, the hybrid score is defined as:

$$\hat{r}_{ui} = \alpha \cdot \hat{r}_{ui}^{CF} + (1-\alpha) \cdot \hat{r}_{ui}^{CB} \tag{2}$$

where the fusion coefficient $\alpha \in [0, 1]$ is dynamically adjusted with the amount of the learner's interaction data — when the cumulative number of interactive resources is lower than the threshold N_{cold} (set to 15 in the experiment), α is 0.2 with content-based filtering as the leading factor; after exceeding the threshold, it is gradually increased to 0.7 by linear interpolation, so that collaborative filtering gradually plays a leading role with the accumulation of data. The cold start problem also presents domain-specific characteristics in the vocational college music performance scenario: freshmen lack behavior data upon enrollment, so the entrance professional evaluation is introduced as the initialization basis, and the three results of intonation score, beat stability and technique completion in the evaluation are mapped to the preset learner prototype clusters. The initial collaborative filtering scores are filled with the average behavior of historical learners within the same cluster who ranked in the top 15% based on their final skill achievement scores and demonstrated a resource completion rate exceeding 90% during their respective semesters. This operational definition ensures that the "high-quality" baseline represents successful learning trajectories that are statistically representative of the current cohort's curriculum standards.

D. CONSTRUCTION FRAMEWORK OF MULTI-DIMENSIONAL PORTRAIT OF VOCATIONAL COLLEGE MUSIC PERFORMANCE LEARNERS

Learner portrait is the information foundation for the recommendation system to perceive individual differences, and the integrity of its construction dimensions directly determines the upper limit of recommendation accuracy. The portrait of vocational college music performance learners needs to cover two levels: static attributes and dynamic behaviors. The static level includes professional direction (vocal music/instrumental music/stage performance), entrance evaluation results, completed courses and certificate records; the dynamic level continuously captures the distribution of practice duration, resource re-access rate, phased skill evaluation scores and teacher feedback labels. After ETL processing, the above data are organized into a structured portrait vector according to the three axes of "professional ability — learning behavior — development intention". The professional ability axis quantifies the current skill achievement rate and shortcoming dimensions; the learning behavior axis depicts learning rhythm, frequency of independent practice and types of resource preference; the development intention axis integrates students' career direction choices and competition registration records to reflect their future performance scenario orientation [6]. The portrait is updated in real time incrementally with the upload of each practice data, and a global recalibration is triggered after each phased evaluation, enabling the recommendation system to perceive the learners' ability leap in different stages and avoid the disconnection between recommended content and actual level due to portrait lag, as shown in Figure 1.

III. SYSTEM ARCHITECTURE DESIGN OF PERSONALIZED TEACHING MODEL

A. COLLECTION DIMENSIONS AND FEATURE ENGINEERING OF PERFORMANCE BEHAVIOR AND VOCAL PRACTICE DATA

The collection of performance behavior and vocal practice data is the raw material supply link for the operation of the personalized recommendation system, and the selection of collection dimensions needs to balance the data availability and the teaching interpretability of features [7,8]. Instrumental music performance data are collected synchronously through MIDI interface and microphone dual channel to obtain note touch force (velocity value, range 0—127), note duration deviation (millisecond-level error relative to the standard score value) and the switching frequency of legato and staccato; vocal practice data rely on mono near-field microphone recording, focusing on extracting the fundamental frequency (F0) trajectory, sound pressure level (SPL) change curve and resonance cavity characteristics. After pre-emphasis, framing and windowing processing of the original audio, the system utilizes a feature selection process based on Information Gain (IG) to identify the most discriminative acoustic properties. For the instrumental direction, a 64-dimensional vector was selected from an initial pool of 128 features (including spectral flux and harmonic-to-noise ratio) as it yielded the highest F1-score in skill-level classification. For vocal music, the 48-dimensional vector was validated against a baseline of standard audio features, demonstrating a 15% higher accuracy in capturing physiological phonation nuances specific to vocational students (including 13-dimensional MFCC, pitch confidence, vocal tract resonance ratio and breath stability index). To reduce the acoustic environment noise caused by different recording devices, mean-variance normalization is applied to all audio features to unify the feature distribution into the zero-mean and unit-variance space. In addition to audio features, the system also collects practice behavior logs synchronously, including single practice duration, number of segment repetitions and distribution of error trigger positions [9]. These behavior metadata are spliced with audio features to form a complete feature sample of the learner's current practice, providing input for subsequent portrait update and recommendation score calculation, as shown in Table 1.

B. STRUCTURED ANNOTATION SYSTEM AND CLASSIFICATION STANDARD OF MUSIC PERFORMANCE TEACHING RESOURCE LIBRARY

The annotation quality of the teaching resource library directly determines whether the content-based filtering algorithm can accurately measure the teaching affinity between resources, so it is necessary to establish a structured annotation system that balances technical interpretability and music professional semantics. The resource library is classified according to two main lines of vocal music and instrumental music, and each main line is divided into three ability levels: basic skill level, intermediate performance level and comprehensive interpretation level. The level division is determined in accordance with the skill achievement indicators in the curriculum standard of vocational college music performance major to avoid annotation inconsistency caused by subjective grading. Each resource record contains seven annotation fields: resource type (etude/demonstration video/theoretical explanation), applicable professional direction, ability level, technical difficulty label (such as passaggio processing/fast running/timbre control), recommended prerequisite resource ID list, average completion duration and average score of historical learners. Technical difficulty labels are managed by a controlled vocabulary, with 28 standard labels for the vocal music direction and 34 standard labels for the instrumental music direction (piano-based). Annotators take up their posts after unified training, and each resource is independently annotated by two annotators, and the annotation consistency rate must reach $\kappa \geq 0.75$ before being put into the library. After being put into the library, the resource generates a 64-dimensional or 48-dimensional content feature vector (corresponding to the feature dimension in Section Collection Dimensions and Feature Engineering of Performance Behavior and Vocal Practice Data) for direct call by the content-based filtering module, and at the same time constructs a directed skill dependency graph with the prerequisite resource ID list as edges, enabling the system to follow the pre-constraints of the skill sequence when generating practice paths [10].

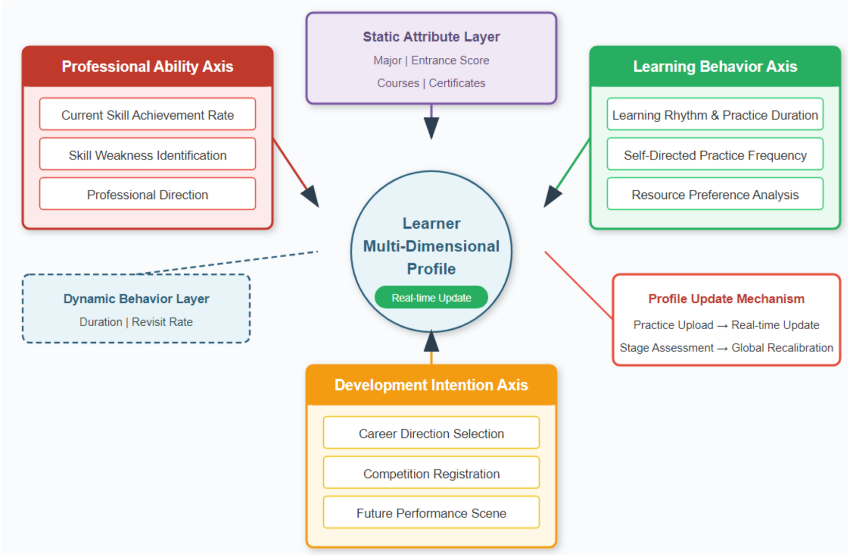

FIGURE 1. Construction Framework of the Three-dimensional and Three-axis Portrait of Vocational College Music Performance Learners

TABLE 1. Collection Dimensions and Feature Indicators of Music Performance Practice Data

Collection Direction	Collection Channel	Core Dimension	Feature Indicator	Feature Dimension / Quantitative Range
Instrumental Music Performance	MIDI Interface + Microphone Dual Channel	Touch Key Feature	Touch key force, note duration deviation, switching frequency of legato and staccato	Force 0–127 / Deviation millisecond-level / Frequency times/minute
Instrumental Music Performance	MIDI Interface + Microphone Dual Channel	Acoustic Feature	MFCC, spectral centroid, roll-off frequency, beat deviation statistics	64-dimensional feature vector
Vocal Music Practice	Mono Near-field Microphone	Physiological Phonation Feature	Fundamental frequency trajectory, sound pressure level change curve, resonance cavity characteristics	F0 value / SPL decibel-level / Resonance ratio
Vocal Music Practice	Mono Near-field Microphone	Acoustic Feature	MFCC, pitch confidence, vocal tract resonance ratio, breath stability index	48-dimensional feature vector
General Practice Behavior	System Background Log	Practice Behavior Feature	Single practice duration, number of segment repetitions, distribution of error trigger positions	Duration minute / Times / Position music score bar number

C. DYNAMIC GENERATION AND REAL-TIME UPDATE MECHANISM OF DIFFERENTIATED PRACTICE PATHS

The generation of differentiated practice paths is essentially a personalized resource sequence planning problem under the constraint of the skill dependency graph. A single static recommendation cannot respond to the dynamic changes of learners' abilities in the practice process, so the system adopts a closed-loop mechanism of "recommendation — execution — feedback — update" to drive the continuous optimization of paths [11]. At the initial stage of path initialization, the system locates the starting node of the learner in the skill dependency graph according to the current skill achievement rate in the learner portrait, carries out a breadth-first search along the directed edges, screens out the top 20 resources with mixed recommendation scores matching the learner's ability level to form a candidate set, and then sorts the candidate set according to the estimated completion duration to generate an initial path covering a two-week teaching cycle. There are two types of conditions for triggering real-time path update: one is that the learner completes the current resource and uploads the practice audio, and the system judges whether the ability has a leap after completing feature extraction and skill scoring; the other is that the learner produces explicit negative feedback (skip or low score) on the recommended resources. The judgment of ability leap takes the skill score improvement exceeding the preset threshold $\Delta s = 0.15$ (full score 1.0) as the trigger condition. After triggering, the system moves the current node to the higher level direction along the dependency graph and calls the recommendation module again to generate the subsequent path segment, making the path difficulty dynamically match the learner's actual progress instead of advancing mechanically according to a fixed plan, as shown in Figure 2.

D. DESIGN OF HUMAN-MACHINE COLLABORATIVE REGULATION INTERFACE WITH THE ACCESS OF TEACHERS' SUBJECTIVE JUDGMENT

Algorithm recommendation performs stably in processing standardized and quantifiable skills, but there are a large number of judgment scenarios in music performance teaching that rely on teachers' professional intuition — such as the protection of students' timbre personality, the evaluation of stage psychological state and the guidance of unconventional performance styles — these dimensions are difficult to be completely captured by feature vectors, so the system designs a collaborative regulation interface for teachers' active intervention. The interface opens three types of operation permissions to teachers: resource locking (forcibly including or excluding a certain resource from the current student's recommended candidate set), path node insertion (inserting teacher-selected resources at the specified position in the

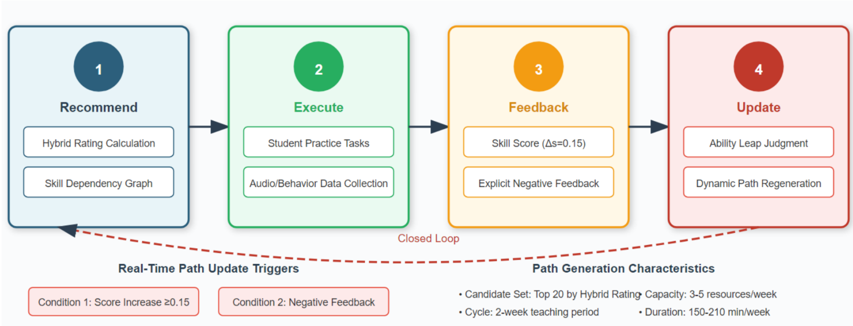

FIGURE 2. Closed-loop Mechanism of "Recommendation - Execution - Feedback - Update" for Personalized Practice Paths

current path) and weight override (manually correcting the score of a specific skill dimension in the student's portrait). Teachers' operation records are stored in the form of event logs with time stamps, and the system transmits teachers' correction behaviors as supervised signals back to the recommendation model. After accumulating enough samples, the model parameters are adjusted through incremental training, enabling the algorithm to gradually acquire the professional judgment preferences of the college's teacher group [12]. To prevent the degradation of personalized effect caused by excessive intervention, the system sets an attenuation window for the influence range of a single teacher's operation: resource locking only acts on the current two-week path segment, and weight override is automatically lifted after the next phased evaluation, ensuring that the main recommendation logic of the algorithm is not replaced for a long time but only revised in stages, and maintaining a reasonable boundary of human-machine collaboration.

IV. EMPIRICAL RESEARCH DESIGN AND SYSTEM IMPLEMENTATION

A. QUASI-EXPERIMENTAL DESIGN SCHEME AND SAMPLE SELECTION

Because the quasi-experimental design cannot randomly assign subjects, it is necessary to control confounding variables in the existing class structure, so the Nonequivalent Control Group Design is adopted to ensure internal validity. The research takes two parallel classes of 2023 grade music performance major in a vocational college as the research objects. To isolate the effect of the personalization algorithm from the impact of digital resource availability, both groups were provided with access to the same digital teaching resource library (3,126 resources) and the practice tracking interface. The experimental group (n=42) accessed the intelligent recommendation engine for personalized path generation, while the control group (n=40) accessed the library via a standard keyword search and a teacher-curated list of suggested materials without algorithmic intervention. A total of 82 students in the two groups all completed a 16-week teaching cycle. Informed consent was obtained from all participants. The pre-test data show that there is no significant difference in the total score of the entrance professional evaluation between the experimental group and the control group (the average score of the experimental group is 67.2 points, the average score of the control group is 66.8 points, independent samples t-test p=0.81), indicating that the basic levels of the two groups are comparable. The two classes are led by the same main course teacher to control for instructional style. To mitigate the confounding effect of the teacher's auditing role, the teacher's interaction with the experimental group was strictly limited to a "passive audit" mode—verifying the technical safety of recommendations without increasing face-to-face instructional time or providing additional qualitative feedback. Furthermore, an independent teaching assistant monitored the control group's platform to ensure equivalent administrative attention was provided to both groups, thus controlling the confounding variable of differential teacher engagement within an acceptable range. The sample size is estimated based on Cohen's d effect size: referring to the medium effect size (d=0.5) of similar educational intervention research reports on recommendation algorithms, under the conditions of statistical power (1-β)=0.80 and two-tailed test with significance level α=0.05, the minimum sample size for each group needs to reach 34, and the actual sample sizes of the two groups all meet this requirement, ensuring the reliability of subsequent statistical tests, as shown in Table 2.

B. DEVELOPMENT AND RELIABILITY AND VALIDITY TEST OF PERFORMANCE SKILL EVALUATION SCALE

The evaluation of music performance skills has long relied on teachers' single subjective scoring, and the evaluation results are significantly affected by the personal style preferences of judges, leading to uncontrollable measurement errors. Therefore, it is necessary to develop a structured evaluation scale with psychometric basis [13]. The scale development has gone through three stages: literature combing, expert interview and pre-test. In accordance with the curriculum standard of vocational college music performance major and Hallam's (2001) theoretical framework on the level of performance skills, an item pool initially formed including 20 scoring indicators in 5 dimensions: intonation and beat, timbre control, technique fluency, phrase expressiveness and overall stage presentation, and each indicator adopts a 5-point Likert scale. In the formal experiment, this scale is used as the main measurement tool for skill

TABLE 2. Comparability Test of Baseline Data between the Experimental Group and the Control Group

Group	Number of People (n)	Average Total Score of Entrance Professional Evaluation (points)	Standard Deviation (SD)	t-value	p-value	95% Confidence Interval	Whether There is a Significant Difference
Experimental Group	42	67.2	5.3	0.25	0.81	[−1.89, 2.69]	No ($p>0.05$)
Control Group	40	66.8	5.6	–	–	–	No ($p>0.05$)

achievement. To eliminate unconscious bias, a double-blind scoring protocol was implemented: performance recordings were anonymized and coded, and the five judges were not informed of the students' group assignments (experimental vs. control) during any of the three evaluation stages. This ensures that the measurement of skill achievement remains objective and free from expectancy bias. The reliability test results show that the overall Cronbach's α coefficient of the scale is 0.90, the α values of the five dimensions range from 0.85 to 0.92, and the inter-rater reliability (ICC) reaches 0.88, all exceeding the psychometric acceptable standards ($\alpha \geq 0.70$, ICC≥ 0.75). The validity test adopts Confirmatory Factor Analysis (CFA), and the fitting indexes of the five-factor model are CFI=0.96 and RMSEA=0.05, the model fitting is good, confirming that the construct validity of the scale meets the requirements [14]. In the formal experiment, this scale is used as the main measurement tool for skill achievement, and three evaluations are conducted before the intervention, at the end of the 8th week and the end of the 16th week respectively, as shown in Table 3.

C. SYSTEM DEPLOYMENT PROCESS AND PHASED IMPLEMENTATION OF TEACHING INTERVENTION

The system deployment is divided into three stages: environment construction, data initialization and formal intervention, with clear and coherent objectives in each stage. In the environment construction stage (Weeks 1—2), the configuration of the server-side recommendation engine and the intranet access authorization of the college are completed, and a unified model of near-field microphone is distributed to each student in the experimental group with the MIDI interface debugging completed to ensure the hardware consistency of the audio collection link [15]. In the data initialization stage (Week 3), the experimental group students are organized to complete the entrance professional evaluation, the evaluation results are automatically parsed by the system and mapped to the preset learner prototype clusters to generate the initial portrait vector of each student, and at the same time, the feature extraction and annotation audit and warehousing of 3126 resources in the teaching resource library (1358 vocal music resources and 1768 instrumental music resources) are completed. In the formal intervention stage (Weeks 4—16), it operates according to the weekly cycle of "system recommendation — student practice — data feedback — path update". Each student in the experimental group receives the personalized practice path pushed by the system every week, the path contains 3—5 resources, and the expected weekly practice duration is controlled at 150—210 minutes to align with the total learning duration of the control group. The mid-term skill evaluation is completed at the end of the 8th week, and the evaluation results trigger the global recalibration of the learner portrait by the system, so that the recommendation strategy in the later 8 weeks is recalibrated on the basis of the mid-term evaluation data, ensuring the continuous adaptability of the intervention intensity throughout the teaching cycle.

D. DATA COLLECTION METHODS AND MULTI-DIMENSIONAL STATISTICAL ANALYSIS STRATEGY

Data collection covers two channels: objective measurement and subjective scale, so as to form a mutually corroborative multi-source evidence chain. Objective data are derived from the system background logs, including the number of resource accesses, the amount of practice audio uploaded, the path completion rate and the recommendation score prediction error (RMSE) of each student in the experimental group; subjective data are collected through two types of standardized scales, among which skill achievement adopts the self-compiled evaluation scale (see Section Development and Reliability and Validity Test of Performance Skill Evaluation Scale), and learning autonomy and learning motivation adopt the two subscales of self-efficacy and intrinsic goal orientation in the Chinese revised version of the Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich et al. (2003) (24 questions in total, Cronbach's α=0.86). The statistical analysis strategy is promoted hierarchically according to the research questions: independent samples t-test is adopted for the comparability test of inter-group baseline; analysis of covariance (ANCOVA) is adopted for the pretest and post-test intervention effects with pre-test scores as covariates to control baseline differences, and the effect size is reported by partial η^2; repeated measures analysis of variance (Repeated Measures ANOVA) is adopted for the longitudinal changes of learning autonomy in the experimental group (three time points: Week 0, 8 and 16); Pearson correlation analysis is adopted for the correlation intensity between the adaptation quality of the recommendation algorithm (measured by RMSE) and the teaching effects of the three dimensions, and Bonferroni correction is applied to multiple comparisons to control the inflation of Type I error, with the significance level uniformly set to α=0.01.

V. ANALYSIS OF EXPERIMENTAL RESULTS AND EVALUATION OF PERSONALIZED TEACHING EFFECTS

TABLE 3. Reliability and Validity Test Results of Music Performance Skill Evaluation Scale

Test Type	Test Indicator	Intonation and Beat	Timbre Control	Technique Fluency	Phrase Expressiveness	Overall Stage Presentation	Overall Scale
Reliability Test	Cronbach's α Coefficient	0.88	0.85	0.92	0.87	0.86	0.90
Reliability Test	Split-half Reliability	0.86	0.83	0.90	0.85	0.84	0.88
Reliability Test	Inter-rater Reliability (ICC)	0.85	0.82	0.89	0.84	0.83	0.88
Validity Test	Confirmatory Factor Analysis CFI	–	–	–	–	–	0.96
Validity Test	Confirmatory Factor Analysis RMSEA	–	–	–	–	–	0.05
Validity Test	Standardized Factor Loading	0.78–0.89	0.75–0.86	0.82–0.93	0.76–0.87	0.75–0.86	0.75–0.93

A. DIFFERENCE TEST OF SKILL ACHIEVEMENT BETWEEN THE EXPERIMENTAL GROUP AND THE CONTROL GROUP

The ANCOVA results, using both the pre-test total skill evaluation score and the total number of resource accesses (practice quantity) as covariates, show that after controlling for these factors, there is still a significant difference in the total skill evaluation score between the groups (F (1,78)=18.52, p<0.001, partial η^2=0.184). This indicates that the improvement in the experimental group is significantly driven by the quality and relevance of the recommended resources rather than merely the volume of practice. The post-test average score of the experimental group is 80.5 points (SD=5.9), and that of the control group is 72.3 points (SD=6.5), with a mean difference of 8.2 points between the two groups, and the experimental group is significantly higher than the control group. The dimensional analysis further reveals the structural characteristics of the differences: in the two technical dimensions of intonation and beat and technique fluency, the average scores of the experimental group are 9.5 points and 8.8 points higher than those of the control group, with extremely significant differences (p<0.001); in the two artistic dimensions of phrase expressiveness and overall stage presentation, the differences between the two groups are relatively narrowed, with mean differences of 5.1 points and 4.5 points respectively, still reaching a significant level (p<0.01). This distribution law indicates that the personalized recommendation system has a more prominent advantage in the quantifiable training of technical skills, while the improvement of artistic expressiveness is more deeply affected by individual temperament and teachers' on-site guidance [16,17], and the intervention effect of algorithm recommendation has a domain boundary. The mid-term evaluation data at the end of the 8th week also show a similar trend (the average score of the experimental group is 75.7 points vs the control group is 70.2 points, p<0.001), indicating that the system effect has begun to appear in the mid-term of the intervention and is further strengthened after the path recalibration in the later 8 weeks, as shown in Figure 3.

B. QUANTITATIVE CHANGE ANALYSIS OF LEARNING AUTONOMY AND LEARNING MOTIVATION

The results of repeated measures analysis of variance show that the experimental group presents a significant main effect of time in the self-efficacy subscale scores (F (2,82)=26.79, p<0.001, partial η^2=0.402), and the average scores at the three time points (Week 0, 8 and 16) are 3.35, 3.82 and 4.31 in turn (full score 5 points), with a cumulative increase of 0.96 points within 16 weeks; the change range of the control group in the same period is only 0.29 points (3.32→3.45→3.61), and the interaction effect of group × time is significant (F (2,158)=18.93, p<0.001), proving that the improvement of autonomy in the experimental group stems from system intervention rather than natural growth effect. The intrinsic goal orientation subscale presents a change trajectory highly consistent with self-efficacy, the average score of the experimental group at the end of the 16th week is 4.25, an increase of 0.87 points compared with the baseline, while the control group only increases by 0.26 points, and the interaction effect is also significant (p<0.001). This result is consistent with the expectation of Self-Determination Theory (Deci & Ryan, 2000): when the difficulty of learning tasks is accurately matched with the individual's ability level, learners' experience of competence is enhanced, which in turn drives the continuous improvement of intrinsic motivation. The personalized recommendation system keeps learners in the "approximate optimal challenge zone" by dynamically adjusting the path difficulty, and at the same time reduces the burnout caused by overly easy tasks and learned helplessness caused by overly difficult tasks, thus realizing the coordinated improvement of learning autonomy and motivation [18,19], as shown in Table 4.

C. EMPIRICAL EVALUATION OF TEACHING RESOURCE MATCHING ACCURACY AND UTILIZATION EFFICIENCY

The system background log data show that the students in the experimental group completed an average of 286 recommended resource accesses within 16 weeks, with a resource completion rate (complete browsing or complete practice) of 82.5%, while the students in the control group completed an average of 165 accesses on the task platform uniformly arranged by teachers, with a completion rate of 57.8%, and the resource utilization efficiency of the experimental group is significantly higher than that of the control group. The matching accuracy of the recommendation system is measured by the Root Mean Square Error (RMSE), and the calculation formula is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{r}_i - r_i)^2} \tag{3}$$

where $\hat{r}_i$ is the system predicted score, r_i is the learner's actual feedback score after completion, and N is the total number of evaluation samples. The RMSE value of the

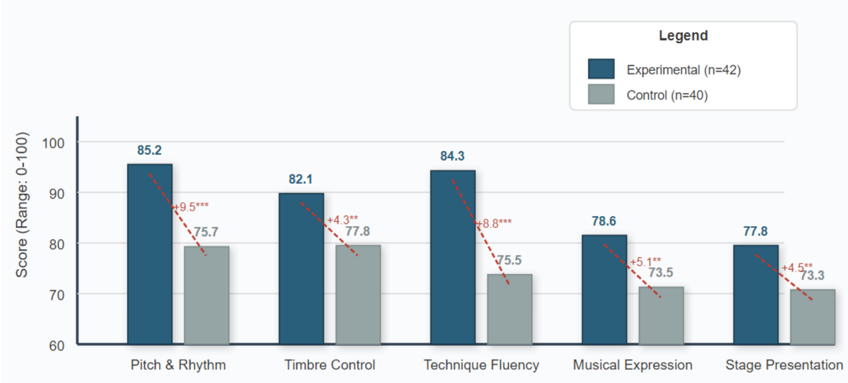

FIGURE 3. Comparison of Post-test Scores of Music Performance Skills in Each Dimension between the Experimental Group and the Control Group

TABLE 4. Longitudinal Scores of Learning Autonomy (Self-efficacy) between the Experimental Group and the Control Group

Group	Score at Week 0 (Baseline) (full score 5 points)	Score at Week 8 (Mid-term) (full score 5 points)	Score at Week 16 (Post-test) (full score 5 points)	Increase at 8 Weeks (points)	Cumulative Increase at 16 Weeks (points)	Increase Rate (16 Weeks)
Experimental Group	3.35±0.42	3.82±0.38	4.31±0.35	0.47	0.96	28.66%
Control Group	3.32±0.45	3.45±0.41	3.61±0.39	0.13	0.29	8.73%
Inter-group Difference (Experimental - Control)	0.03	0.37	0.70	0.34	0.67	19.93%

system shows a continuous downward trend during the intervention: it is 0.78 at the end of the 4th week, drops to 0.56 at the end of the 8th week, and further drops to 0.42 at the end of the 16th week, indicating that with the accumulation of learners' behavior data, the personalized fitting ability of the hybrid recommendation model is significantly enhanced. The matching quality can be further verified by comparing the resource skip rate: the skip rate of recommended resources is 23.6% in the 4th week of the intervention, and drops to 8.9% by the 16th week, with a drop of 62.3%. This trend shows that after the cold start period, the fit between the system's recommended content and the learners' actual needs is continuously improved, and the learners' acceptance of the recommendation results is enhanced accordingly, thus forming a positive feedback loop of "data increase → recommendation accuracy improvement → learners' acceptance improvement → further data increase".

D. COMPREHENSIVE IMPACT OF RECOMMENDATION ALGORITHM ADAPTATION QUALITY ON THE THREE-DIMENSIONAL TEACHING EFFECTS

To test the systematic correlation between the adaptation quality of the recommendation algorithm and the three-dimensional teaching effects, each student's final RMSE (denoting recommendation precision) was utilized as the predictive variable for Pearson correlation analysis at the end of the 16th week, and Pearson correlation analysis is conducted with the skill achievement score, learning autonomy score and resource completion rate respectively. The refined algorithm improves recommendation accuracy for unstructured artistic data through domain-specific adaptive transformation. After Bonferroni correction, the significance threshold is adjusted to α=0.017. The results show that RMSE is significantly negatively correlated with skill achievement ($r=-0.65$, $p<0.001$), significantly negatively correlated with learning autonomy score ($r=-0.58$, $p<0.001$), and also significantly negatively correlated with resource completion rate ($r=-0.62$, $p<0.001$), and all three groups of correlations remain statistically significant after correction. The above results indicate that the higher the adaptation quality of the algorithm for individual learners, the better their performance in the three dimensions of teaching effects, and the three paths are not independent but mutually reinforcing: high-adaptation recommendation reduces the proportion of ineffective practice, improves the skill gain per unit time, further enhances learners' experience of competence, and leads to a further improvement in learning autonomy [20]. It is worth noting that the absolute values of the correlation coefficients do not exceed 0.70, indicating that the adaptation quality of recommendation is not the only factor determining the teaching effect. The differences in teachers' intervention quality, students' individual effort and practice conditions together constitute the regulatory boundary of the system effect, which further confirms the necessity of human-machine collaborative design in the personalized teaching system.

VI. CONCLUSION

The empirical results confirm that the personalized teaching model driven by intelligent recommendation algorithms is effectively applicable to the music performance major in vocational colleges. The algorithm after adaptive transformation effectively improves the recommendation accuracy of unstructured art data and can maintain stable recommendation quality even when data is sparse in the early stage of teaching. The differentiated practice paths have significantly improved students' skill achievement and learning autonomy, realizing the transformation from algorithm empowerment to teaching effectiveness. The skill evaluation scale tested for reliability and validity has effectively reduced the subjectivity of art evaluation; the human-machine collaboration mechanism has balanced algorithm efficiency and teachers' professional judgment, enhancing practical applicability. The research clarifies that the depth of algorithm domain adaptation, the accuracy of evaluation tools and the systematization level of teaching intervention are the three key factors to improve the effectiveness of personalized teaching, which provides empirical basis for the optimization of intelligent teaching systems for art majors in vocational colleges.

AUTHOR CONTRIBUTIONS

Conceptualization – Yixin Cao, Na Li, Chenlu Liu, Ping Wang, Nina Gao, Jiangbo Guo and Wen Song; methodology – Yixin Cao, Na Li, Chenlu Liu, Nina Gao, Jiangbo Guo and Wen Song; investigation – Yixin Cao, Na Li, Chenlu Liu, Ping Wang, Nina Gao, Jiangbo Guo and Wen Song; writing-original draft preparation – Yixin Cao, Na Li, Chenlu Liu, Ping Wang, Nina Gao, Jiangbo Guo and Wen Song. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Y. Wang, Q. Yang, and L. Jiang, "Innovative Practice of "Performance-Creation Integration" Teaching Model for Music Majors in Higher Education," New Explorations in Education and Teaching, vol. 3, no. 12, 2025, DOI: 10.70711/NEET.V3I12.8261.

[2] Y. Z. Abby and H. Jing, "Collaborative Filtering Recommendation Technology for College Music Teaching Strategy and Security Application Under Deep Learning," International Journal of Information Technologies and Systems Approach (IJITSA), vol. 18, no. 1, pp. 1-18, 2025, DOI: 10.4018/IJITSA.393670.

[3] S. Guo, "Design of Vocal Music Performance Teaching System Based on Multimedia Intelligent Platform," International Journal of High Speed Electronics and Systems, vol. 35, no. 03, 2025, DOI: 10.1142/S0129156425404152.

[4] N. Li and Y. Xiao, "Design of an Intelligent Teaching Platform under Multidimensional Data Fusion in Music Performance Teaching," Applied Mathematics and Nonlinear Sciences, vol. 10, no. 1, 2025, DOI: 10.2478/AMNS-2025-0807.

[5] H. Wang, "Analysis of Teaching Mode of Music Major Students Based on Personalized Recommendation Algorithm," International Journal of High Speed Electronics and Systems, to be published, 2024, DOI: 10.1142/S0129156425401275.

[6] B. Wang and P. Li, "Personalized recommendation based on improved speech recognition algorithm in music e-learning course simulation," Entertainment Computing, vol. 52, 100721, 2025, DOI: 10.1016/J.ENTCOM.2024.100721.

[7] K. Li, "Innovative Strategies in Teaching Vocal Music for Higher Vocational Music Performance Majors," Art and Performance Letters, vol. 5, no. 2, 2024, DOI: 10.23977/ARTPL.2024.050222.

[8] Y. Wu and W. Kang, "Research on the Application of Music Technology in Modern Music Education in the Information Age," Applied Mathematics and Nonlinear Sciences, vol. 9, no. 1, 2024, DOI: 10.2478/AMNS-2024-1147.

[9] L. Liu, "Big data analysis for sustainable music teaching using collaborative filtering recommendation optimisation algorithm," International Journal of Computational Systems Engineering, vol. 8, no. 3-4, pp. 273-282, 2024, DOI: 10.1504/IJCSYSE.2024.142772.

[10] G. Hongjiao and L. Fan, "The application of virtual reality technology in the teaching of clarinet music art under the mobile wireless network learning environment," Entertainment Computing, pp. 49, 2024, DOI: 10.1016/J.ENTCOM.2023.100619.

[11] J. Li, "Music Performance Major "Internet +" Teaching Mode Application Strategy," International Journal of Mathematics and Systems Science, vol. 6, no. 3, 2023, DOI: 10.24294/IJMSS.V6I3.2659.

[12] P. Richard, "Sound teaching: a research-informed approach to inspiring confidence, skill, and enjoyment in music performance," Music Education Research, vol. 25, no. 1, pp. 118-120, 2023, DOI: 10.1080/14613808.2022.2146355.

[13] Peng C .A Personalized Recommendation Method for Vocal Music Teaching Resources Based on Network Topology.International Journal of High Speed Electronics and Systems, 2026, 35(03), DOI: 10.1142/S012915642540467X.

[14] Liu D , Zou Y .Innovation and Practical Exploration of Music Classroom Teaching Models Assisted by Digital Technology.International Journal for Housing Science & Its Applications, 2025, 47(1), DOI: 10.70517/ijhsa47114.

[15] E, M. G, B. Jennifer, and H. John, "Feedback in Music Performance Teaching," Frontiers in Psychology, vol. 13, 891025, 2022, DOI: 10.3389/FPSYG.2022.891025.

[16] T. Changfei and Z. Jun, "An intelligent deep learning-enabled recommendation algorithm for teaching music students," Soft Computing, vol. 26, no. 20, pp. 10591-10598, 2022, DOI: 10.1007/S00500-021-06709-X.

[17] H. Meissner, R. Timmers, and S. Pitts E, "Sound Teaching: A Research-Informed Approach to Inspiring Confidence, Skill, and Enjoyment in Music Performance," Taylor and Francis, 2021, DOI: 10.4324/9781003108382.

[18] YanY, "Strategies to Effectively Dock Professional Music Performance Skills Teaching and Stage Practice," Journal of Contemporary Educational Research, vol. 5, no. 11, pp. 1-4, 2021, DOI: 10.26689/JCER.V5I11.2695.

[19] C. Wei, "Research on School-enterprise Cooperation Training Model of Music Performance Professionals in Colleges and Universities Based on Repertoire Creation and Computer Multimedia Technology," Journal of Physics: Conference Series, 1992(2), 2021, DOI: 10.1088/1742-6596/1992/2/022033.

[20] V. Malaschenko O, M. Antonova A, G. Knyazeva L, et al., "Theoretical and Practical Aspects of Integrating the Computer and Music Technologies into The Instrumental Performance Training of The Pedagogical University Students," SHS Web of Conferences, pp. 79, 2020, DOI: 10.1051/shsconf/20207901014.