ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Role Reconstruction and Professional Development Path of Music Teachers in the Age of Artificial Intelligence

Fen Pi[1] *and Wei Zheng*[2,*]

[1] Kashi University, Kashi, Xinjiang, 844000, China,Zhaoqing University, Zhaoqing, Guangdong, 526061, China
[2] Linyi University, Linyi, Shandong, 276000, China

SUMMARY: *From the perspective of how to understand students better, the study provides a feasible path for teachers' differentiated teaching and role reconstruction by constructing learner profiles and clustering and subclustering. The "learning-learner" feature layer is constructed, and the multimodal data of learners are integrated from six dimensions, such as intrinsic layer, learning layer, and expectation layer, and transformed into six dimensions, such as theory grade, performance grade, cooperation frequency, independent practice, and so on. It is transformed into seven quantifiable music learning effect indicators, such as theory score, performance score, cooperation frequency, and independent practice time. Canopy-Kmeans clustering algorithm was used to identify different types of students, and real data from 124 music performance majors were selected for analysis. The students' overall theoretical and practical scores amounted to 83.02 and 76.75, and the average expected match was 89.83%, which was a good overall performance but with significant individual differences. Correlation analysis further revealed that peer evaluation and expectation fulfillment were highly correlated with r=0.919. Through Canopy-Kmeans clustering, the students were clearly classified into three categories, with 25 students of the leading type who performed comprehensively and were significantly excellent in all the indicators, especially in theory (95.8), practice (93.24), and frequency of collaboration (52.84 times). The 85 students of the intermediate type were in the middle to upper range of all indicators and had the greatest improvement in skills (2.67). In contrast, 14 of the lagging type were significantly lower in theory, practice, practice engagement and collaborative participation.*

KEYWORDS: *music teacher; Canopy-Kmeans; learner profiling; role transformation*

1 Introduction

In recent years, the breakthrough progress in the field of artificial intelligence (AI) has had a profound impact on the field of education, and the impact of AI on education is mainly carried out through two ways: first, the industrial revolution and work revolution brought by AI has triggered a shift in the goal of education, forcing education to change; and second, AI has changed the concept of education and the way of education, prompting the education reform to develop in depth [1-3]. As far as education is concerned, AI is undoubtedly a double-edged sword, which impacts teachers' education and teaching while giving them a boost.

In the era of AI, the internal mechanism of teacher professional development is more sophisticated, and the system concept of teacher professional development presents an

*18818871668@163.com
https://doi.org/10.65102/is2026263 Published:30 April, 2026

unprecedented new look. The new generation of information technology such as the Internet, AI, and big data has also brought greater opportunities for teacher professional development: on the one hand, teachers are adapting to the changes in the surrounding environment and their professional roles. On the other hand, the teacher professional development system has changed from formal changes to substantive changes, forcing teachers to seek a broader and more sustainable time and space for professional development [4-7]. However, at the same time, the development of the AI era has brought more and more complex challenges to teachers' professional development, challenging their authoritative position, deconstructing their professional status, increasing the difficulty of student management, and impacting their role position and development in education [8-11]. Based on this, teachers need to constantly adjust their own positioning and professional development, and reconstruct the role of teachers in order to do a good job of leading the development of students in the new era.

In the era of AI, the field of music education is also facing profound changes, from the new paradigm of music creation, personalized learning, tract music resource allocation and virtual teaching, etc., which address the challenges of the traditional music teaching model [12-14]. It is precisely for this reason that music teachers need to re-examine their own value, adjust their own positioning and professional development, enhance their professional competitiveness and career development ability, enhance the cultivation of intelligent literacy, improve the ability of intelligent teaching, and move from experience teaching to intelligent teaching [15, 16]. Only by changing concepts, strengthening learning, adjusting role positioning, and focusing on professional development can music teachers adapt to the changes of the AI era and promote the development of music education.

To respond to the topic of how music teachers are technologically empowered to achieve professional development in the age of artificial intelligence, a return to the student body is needed. The study makes teaching interventions more personalized by drawing exclusive learning portraits for students and clustering to find learning groups with similar characteristics. Firstly, we build a “learning-learner” feature layer framework to capture students' learning behaviors, interaction heat, mood swings (learning layer), as well as their goals, styles, and environments. It is then translated into seven quantifiable learning outcome indicators, including theoretical mastery, practical performance, autonomy, cooperation, etc., and assigned with weights. The clustering method integrating Canopy and K-means is further adopted. Based on the students' performance on each index, they are categorized into different clusters. Finally the real application is applied to music learning data in colleges and universities. From performance collection, grade conversion, to feature extraction and clustering modeling, it helps music teachers to adjust teaching strategies for different groups.

2 Construction and analysis of music learner portrait based on multidimensional features and clustering

2.1 Learner Digital Portrait Construction and Data Acquisition

2.1.1 Construction of a digital portrait based on the “learning-learner” feature layer

The “learning-learner” feature layer is at the theoretical level, integrating learning and learner-related multimodal data such as learning logs, psychology, physiology, etc., and characterizing learning activities and learner attributes through qualitative analysis [17]. Learning analytics technology allows the portrayal of learner user profiles to be transformed from a static point construction to a dynamic and continuous construction, and the emergence of e-learning platforms has enhanced the clarity of learner user profiles. Compared to the traditional construction of learner user profiles with learning imprint data, the addition of learner performance assessment and self-reporting data can reduce the error rate of learner user profiles in predicting learning performance by four to eleven percentage points, and provide an opportunity for In order to enhance the diversity and saturation of learner data, this study constructs a “learning-learner” feature layer and divides it into six sub-layers: intrinsic layer, attribute layer, learning layer, expectation layer, evaluation layer, and environment layer, and the relationship of each layer is shown in Figure 1.

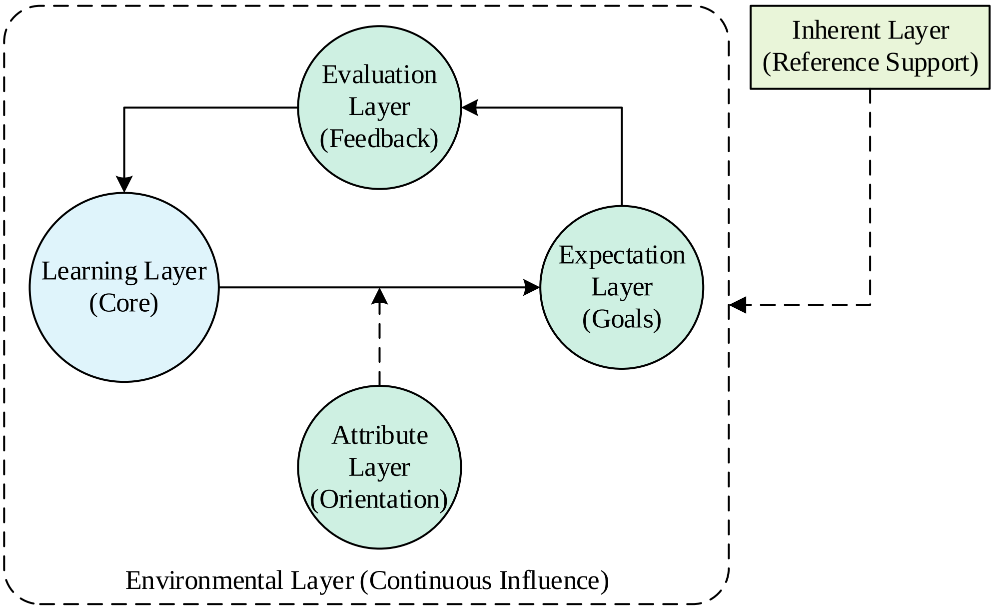


Figure 1: "Learning - Learner" Feature Layer

The intrinsic layer plays an auxiliary role in the construction of the portrait, and the data in the layer include learners' demographic characteristics (gender, age, etc.) and sociological characteristics (occupation, title, profession, etc.), and the data in the intrinsic layer are mostly static data, which can be obtained by the researcher through questionnaires at one time. The learning layer is the core of portrait construction, including learning behavior, communication and interaction, emotional changes, physiological fluctuations, and a series of naturally occurring data generated by the learner during the learning process, which is the specific performance of the time frame, tool frame, and object frame in the frame centered on “learning”.

The expectation layer is the learner's learning goals, including the learner's expectations, the teacher's expectations, and the school's expectations. School expectation is a prerequisite for school management to formulate development goals for teachers and students according to the direction and concept of school development; teachers, as the conveyor of school expectation, should be driven by school expectation to put forward corresponding requirements for students' learning activities; and learners should formulate corresponding goals according to their own actual learning situation. The attribute layer plays a guiding role and is a reflection of learners' characteristics and attributes, including data on learners' learning styles, living habits, and personalities. Based on the data in the attribute layer, the researcher develops a personalized learning plan for the learner and continuously guides the learner to achieve the desired goals.

The evaluation layer plays the role of feedback, focusing on the entire learning process of formative evaluation as a means of self-evaluation, peer evaluation, teacher evaluation and systematic evaluation to make learning outcomes three-dimensional, and timely adjustment of the learner's personalized learning program. The environment layer mainly refers to the learner's learning environment, which has a continuous impact on the learner's learning behavior. In today's intelligent era, the learning environment is diverse and complex, so it fully understands and makes use of the characteristics of the field teaching environment, the intelligent learning environment, the ubiquitous learning environment, the virtual learning environment, and collects data on the environment's temperature, humidity, lighting, and the comfort level of the equipment used continuously.

2.1.2 Identification and quantification of learning effectiveness indicators

Learners' music learning data and indicators are synthesized into seven indicators according to their categories, which are established and quantified to provide a data basis for the subsequent classification of learner types. The seven indicators reflect learners' learning effects, namely, the degree of theoretical learning, the degree of practical learning, the degree of independent learning, the degree of interaction and communication, the level of peer evaluation, the degree of improvement in competence, and the value of the total expectation achievement. Among them, the degree of theoretical learning, the degree of practical learning and the degree of interactive communication correspond to the learning layer of the “learning-learner” characteristic layer; the level of peer evaluation corresponds to the evaluation layer, the degree of independent learning corresponds to the attribute layer, and the degree of competence enhancement and the degree of expectation fulfillment correspond to the expectation layer.

The four indicators, i.e. theoretical learning level, practical learning level, interactive communication level, and independent learning level, reflect the size of learners' learning ability; the peer evaluation level is reflected by the online mutual evaluation scores; the level of competence enhancement reflects the change of relevant competence before and after the course, which is obtained by organizing and quantifying relevant questions in the pre-learning questionnaire and self-reporting questionnaire filled out by learners before the course and after the course. Learners' expectations are divided into expectations for the course content and expectations for the teachers, which to a certain extent can reflect the learners' satisfaction with the course and have a certain guiding significance for the course improvement, and are obtained by setting up relevant topics in the questionnaires before and after the course.

For the five indicators of learners' theoretical learning degree, practical learning degree, independent learning degree, interactive communication degree and peer evaluation level, the hierarchical analysis method was used to construct a three-level analytical structure model, with the first level as the goal, the second level as the data acquisition method, and the third level as the specific problem. The binary comparison method was used to establish the relative importance of the data acquisition methods in the second level, and the corresponding weights were assigned to them through practical experience; for the allocation of weights to specific issues in the third level, the coefficient of variation method was used to calculate the weights of each issue, as shown in Fig. 2.

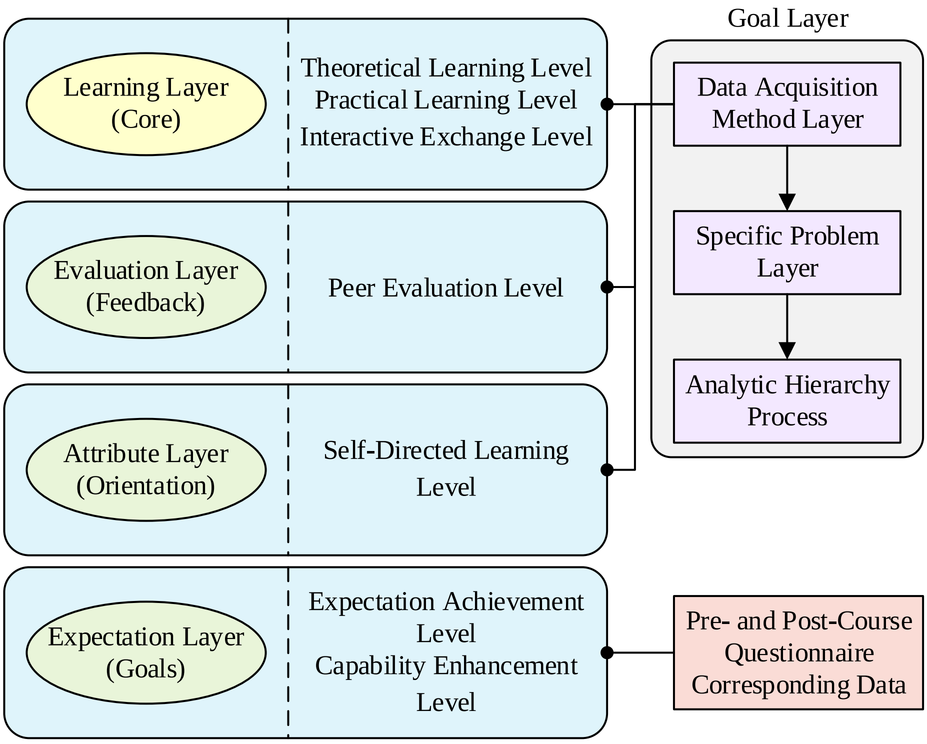


Figure 2: Quantitative Method for Learning Effect Indicators of "Learning - Learner"

The coefficient of variation is an important indicator for portraying the tendency of data to deviate from the central value, reflecting the difference and fluctuation of the values taken, which is numerically equal to the standard deviation divided by the mean value. The use of the coefficient of variation to determine the weight of factors in the evaluation system is a concise and effective means, in the evaluation system, if a large difference in the value of a factor, it means that the factor is difficult to achieve, is a key factor in the response to the gap in the object under evaluation, and such factors should be given a higher weight.

In the process of practical application, then there is:

$$W_i = \frac{v_i}{\sum_{i=1}^{n} v_i} \tag{1}$$

where V_i - coefficient of variation of the i th specific question under a layer of the data acquisition layer;

W_i - the weight coefficient of the i th specific problem under a certain layer of the data acquisition layer.

So far, the method of assigning weights to the two levels of the data acquisition method layer and the specific problem layer has been established, and through brief calculation, the quantitative data of the target layer can be obtained, and the quantization of the five indicators can be completed, which lays the data foundation for the subsequent classification of learner types.

2.2 Clustering Algorithm

After determining the quantitative indicators to be studied in this paper, in order to identify the types of learners with similar characteristics from the above student music data, the Canopy-

Kmeans clustering algorithm is now introduced to realize the music learner typing by pattern recognition of students' performance on each indicator.

2.2.1 Canopy clustering algorithm

Canopy algorithm is an unsupervised preclustering algorithm often used as a preprocessing step for K-means algorithm or hierarchical clustering algorithm [18]. Clustering algorithms are categorized into coarse and fine clustering, and the Canopy clustering algorithm is a coarse clustering algorithm. Unlike the fine clustering algorithm, Canopy clustering algorithm does not require the implementation of a specified number of clusters, the execution speed is very fast, suitable for dealing with large-scale datasets, but there is no way to guarantee the accuracy of the clusters.

Canopy algorithm for a given dataset, you need to pre-set two distance thresholds R_1 and R_2, randomly select the data point as the initial clustering center, and calculate the Euclidean distance of the sample objects in the dataset to the initial clustering center, through the distance thresholds to determine the subordinate relationship of the sample objects, until all the sample objects are divided up, and ultimately get the number of class clusters and initial clustering center. Canopy algorithm is shown in Figure 3.

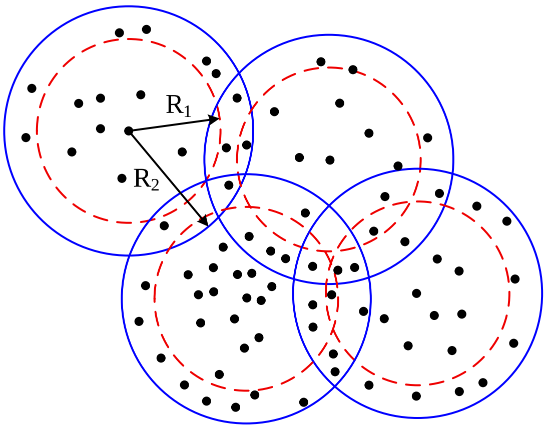


Figure 3: Canopy clustering diagram

The steps of Canopy algorithm are as follows:

(1) Input the dataset $X = \{x_1, x_2, \cdots, x_n\}$, and pre-set the distance thresholds R_1 and $R_2 \left(R_2 > R_1\right)$ through the dataset.

(2) Randomly select a point x_i from the dataset X, use x_i as the clustering center of the first cluster, and remove the point from X.

(3) Continue to randomly select a point x_j from the remaining dataset X, calculate the distance d from x_j to all the generated cluster centers, and if the distance $d < R_1$ to a particular cluster center, then add the point x_j to that cluster, and remove the point x_j from the dataset X and x_j is not added to any other clusters (this point is similar enough to the cluster center) is sufficiently similar to no longer serve as other cluster centers). If the distance $d > R_2$ from x_j to all cluster centers, use x_j as a new cluster center and also remove the point from the dataset X. If the distance $R_2 \geq d \geq R_1$ add the point P to the cluster but do not remove the point from the dataset X (the point continues to participate in the calculation).

(4) Repeat step (3) until the data set X is null.

2.2.2 K-Means clustering algorithm

K-Means in the data clustering, according to the data set itself has different data attributes, according to certain rules to calculate the similarity between the data, and finally, the data will be divided into different classifications, for the typical machine learning algorithms, which are widely used. Figure 4 shows the K-Means algorithm clustering execution process.

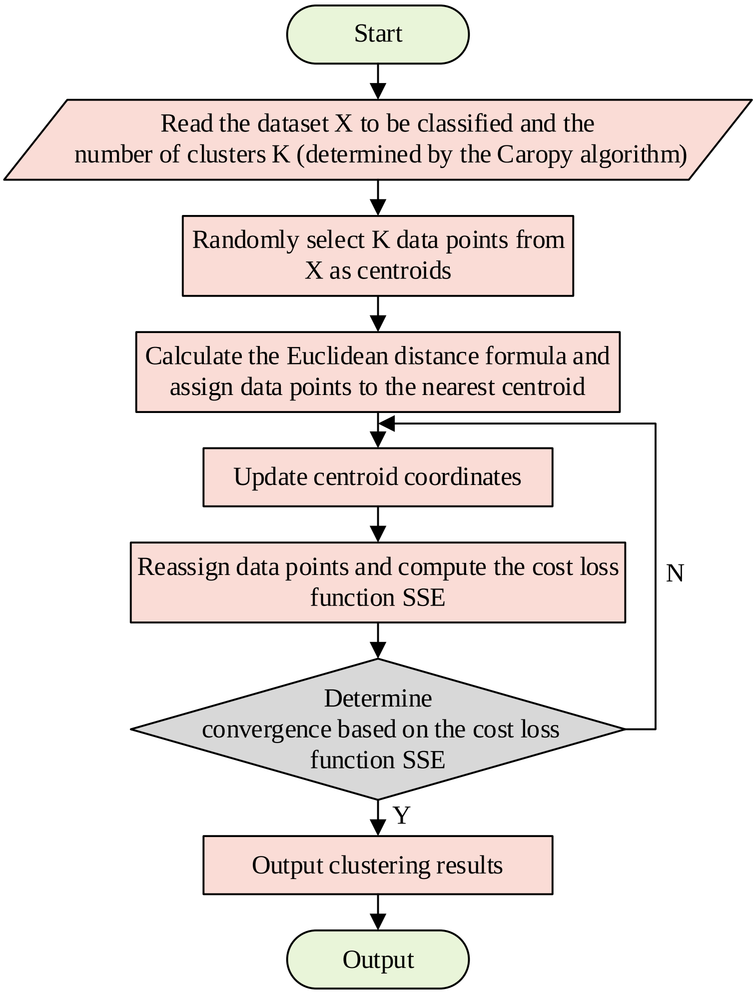


Figure 4: The execution process of the K-Means clustering algorithm

The clustering process of the K-Means algorithm, where X denotes the dataset to be categorized, can be defined in the form of Eq. (2), which denotes that X has n data points and each data point has M attribute values. In Fig. 4, the initial center of mass is selected by the Canopy algorithm to determine the k value, whose matrix representation is shown in Eq. (3) and Eq. (4). The Euclidean distance calculation formula used in Fig. 4, denoted as EUD, is expressed as Eq. (5), and the data point X_n is attributed to the class in which the nearest distance C_k is located, as shown in Eq. (6). The minimization SSE function used, mathematically represented in equation (7). j denotes the order of the elements in the set, μ_k denotes the overall set of elements to which it belongs and to which it belongs to the class k, and C_i denotes the center of mass and to which it belongs to the class i.

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \cdots & x_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & x_{mn} \end{bmatrix} \tag{2}$$

$$InitialCenteroid = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1k} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2k} \\ c_{31} & c_{32} & c_{33} & \cdots & c_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & c_{m3} & \cdots & c_{mk} \end{bmatrix} \tag{3}$$
$$\left(k \in 1,2,3,\cdots,K,\ m \in 1,2,3,\cdots,M\right)$$

$$C_k = \begin{bmatrix} c_{1k} \\ c_{2k} \\ c_{3k} \\ \vdots \\ c_{mk} \end{bmatrix} \tag{4}$$
$$\left(k \in 1,2,3,\cdots,K,\ m \in 1,2,3,\cdots,M\right)$$

$$EUD\left(X_n, C_k\right) = \sqrt{\sum_{m=1}^{m}\left(x_{mn} - c_{mk}\right)^2} \tag{5}$$
$$\left(n \in 1,2,3,\cdots,N,\ k \in 1,2,3,\cdots,K\right)$$

$$X_i = \left\{Min\left(EUD\left(X_n, C_k\right)\right)\right\} \tag{6}$$
$$\left(i \in \left\{1,2,3,\cdots,N\right\},\ n \in 1,2,3,\cdots,N\right)$$

$$SSE = \sum_{i=1}^{k}\sum_{j\in\mu_k}\left(x^{(j)} - c_i\right)^2 \tag{7}$$

2.2.3 Canopy-Kmeans Clustering Algorithm

Canopy-Kmeans is an improved algorithm for K-means and Canopy [19], whose central idea is to use Canopy to preprocess the dataset for coarse clustering first, quickly divide the data into each Canopy set by distance, and then use the Canopy centroid as the initial center of mass point of the K-means algorithm for K-means Clustering.

The Canopy-Kmeans algorithm flow is shown in Fig. 5, and the specific steps are as follows:

(1) The original data set R is sorted according to certain rules, and the two thresholds $D1$ and $D2$ of Canopy are set, where $D1 > D2$.

(2) Pick a data vector i randomly from the set R, put the vector i into the Canopy set (initially the empty set) as a Canopy center point, and remove i from the set R.

(3) Calculate the distance between i and other data samples, and determine whether other vectors in R belong to the current Canopy. firstly, find the distance d_{ij} between the current vector $j\left(j \neq i\right)$ and i, if $d_{ij} > D1$, it means that the vector j does not belong to the current Canopy; if $D2 < d_{ij} < D1$, the vector j is added to the current set of Canopy; if $d_{ij} < D2$, the vector j is added to the current set of Canopy and j is removed from R. Loop steps (2) and (3) until R becomes the empty set.

(4) Use the centers of the k Canopy obtained from the above steps as the clustering centers.

(5) Find the distance between each sample point and the cluster center in each Canopy class, and classify the point into the cluster with the smallest distance.

(6) In each cluster, calculate the average value of the cluster, and find the point closest to the average value as the new clustering center.

(7) Determine whether the results of clustering converge, that is, whether the cluster center is shifted or not, if it converges then stop, if it does not converge then loop steps (5) and (6) to continue until convergence.

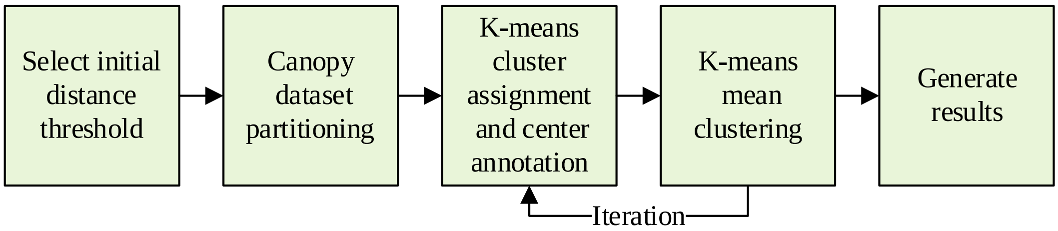


Figure 5: Canopy-Kmeans algorithm flow

Canopy-Kmeans algorithm needs to set the initial distance thresholds $D1$ and $D2$ in practical applications, these two values are generally set by human beings in practical applications, such a selection method has a certain degree of uncertainty, improper selection may lead to the final clustering results are not ideal.

2.3 Analysis of college students' music learning performance

2.3.1 Sources of data on academic performance in music

The music learning performance of university students is an important indicator of the effectiveness of their music learning and the quality of their teaching. The performance data generated from the school's related music course exams, skills assessment, performance practice and other activities is an important basis for analyzing students' music learning performance. It not only reflects the teaching level of teachers, but also allows students to understand their own deficiencies and strengths, thus helping teachers and students to find the problem in time and adjust the teaching program and learning methods.

With the increase in the number of students enrolled in colleges and universities, it is not only time-consuming and laborious to conduct simple statistics on students' performance in the past, but also cannot get useful feedback information. Therefore, this paper adopts the above-mentioned Canopy-Kmeans algorithm to analyze the seven music learning effect indicators obtained in section 2.3.2, to accurately locate each student, and at the same time, it can obtain the strengths and weaknesses of each type of student in vocal performance, which provides music teachers with the basis for developing teaching methods, and then achieves the purpose of improving the quality of teaching.

2.3.2 Data pre-processing

(1) Data selection

In order to ensure the authenticity and rationality of the data, we need to complete the selection and organization of the original data. In this paper, we select the above seven music scores of 124 students in the 2024 session of a college majoring in music performance as the analysis data, and each student's score is stored in a table in the form of percentage.

(2) Data integration

Data integration is to centralize and merge files or data from different data sources together to facilitate data analysis and processing. The main work of this process is to solve the contradictory problems of multiple data sources, such as the inconsistency of word meaning, structure, name, etc. We need to unify and standardize the processing of these data to form the initial data suitable for cluster analysis. In this paper, the student's academic number is used as the primary key to process the music performance, and the collection of music performance of each student is organized into a performance table.

(3) Data Cleaning

Data cleaning mainly includes removing redundancy, noise and other abnormal data, processing vacant values (deleting or filling) and data format conversion. For the vacant values in the grade information table, if they are caused by disciplinary reasons or absence from the examination, they are uniformly recorded as zero.

For the vacant values in the grade information table, if they are caused by disciplinary reasons or absence from the examination, then they are uniformly recorded as zero and deleted; if all the grades of a student are empty, it is considered that the student's academic performance has changed.

If all the grades of a student are null, it will be considered that the student's academic performance has changed and all the data information of the student will be deleted.

If the omission is caused by the teacher's entry of information, the average grade will be used to fill in at this time.

(4) Data conversion

After cleaning and organizing the results of data conversion, mainly for peer assessment “excellent, good, medium, poor” and other forms of records into {4,3,2,1} scores.

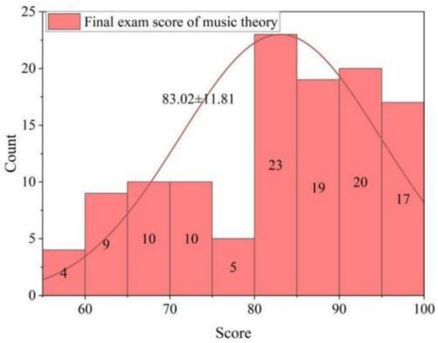
3 Cluster analysis and feature discovery of music learning groups

Based on the multi-source music learning data of 124 music performance majors, this chapter firstly grasps the intrinsic connection between the overall performance of the students and various indicators through descriptive statistics and correlation analysis; then, it applies Canopy-Kmeans clustering method to objectively group the students; finally, it focuses on the performance characteristics of the different groups to reveal their learning patterns and potential needs, which provides a data basis for the subsequent exploration of the teachers' role Finally, we focus on the performance characteristics of different groups to reveal their learning patterns and potential needs, providing data basis for the subsequent discussion of teacher role transformation.

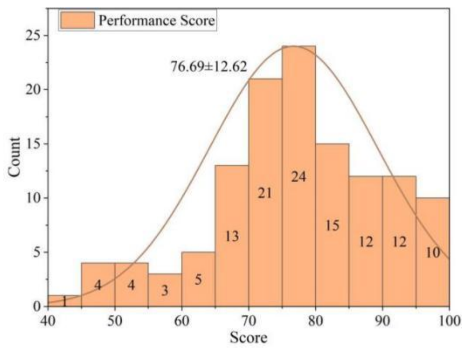
3.1 Analysis of learning outcomes

3.1.1 Descriptive statistical analysis of the sample

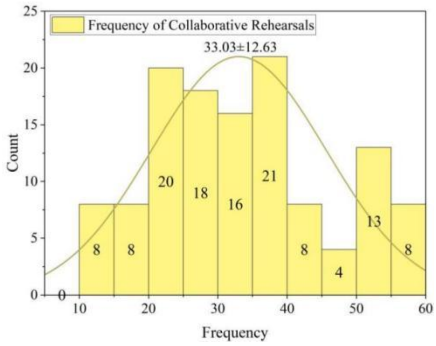
The final grade of music theory, singing/playing scores, frequency of collaborative rehearsals, weekly hours of independent practice outside the classroom, peer evaluation, comparison of music skill tests before and after the course, and the degree of matching between students' course expectations and final achievements were used to refract the students' degree of theoretical learning, level of practical learning, level of independent learning, level of interaction, level of peer evaluation, degree of competence enhancement, and the value of total expectation achievement. Total Expectation Realization Value. Firstly, descriptive statistics were analyzed on the seven music-related scores collected from 124 music performance majors, and the results are shown in Figure 6.



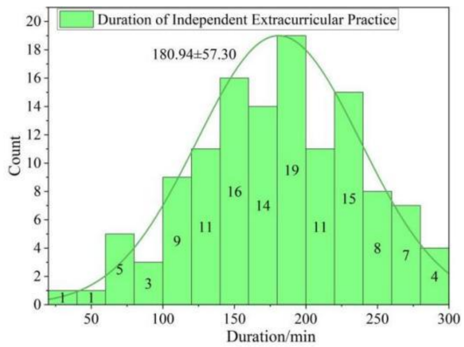
(a)Final grade of music theory



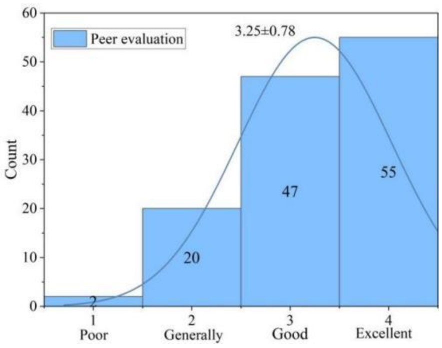
(b)Performance score



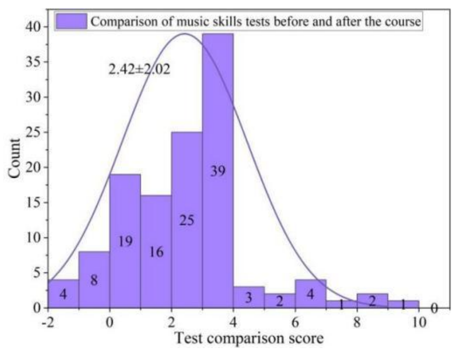
(c)Frequency of collaborative rehearsals



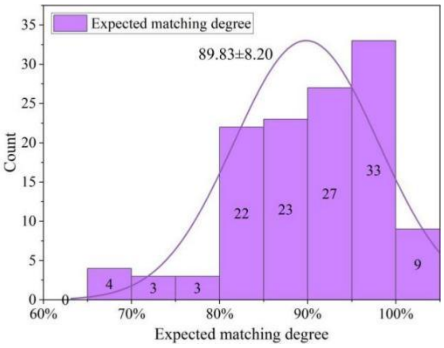
(d)Duration of independent extracurricular practice



(e)Peer evaluation



(f)Comparison of music skills tests before and after the course



(g)Expected matching degree

Figure 6: Descriptive statistical analysis of seven music-related performance indicators

As a whole, students performed solidly in their music theory final grades, with a mean score of 83.02±11.81, and a median score of 85. 37 students, or 29.84%, scored 90 or more, indicating that the majority of students' scores were concentrated at the middle to high level. There were 4 students in the failing band of <60 points. There are still differences in individual mastery of music theory, and not all students mastered it easily.

In contrast, singing/playing scores were more varied, with an average of 76.75±12.70 points and a median of 77 points, while the 70-80 intermediate band accounted for the largest proportion of students, at 45/124=36.29%; there were only 22 students in the >90 high band, and the lowest score in playing was even only 42, reflecting the polarization of students in this category in the subject of music performance.

In terms of the frequency of collaborative rehearsals, which reflects the degree of interaction and communication, students on average participated in a total of about 33 times a semester, but the span ranged from 10 to 58 times. The majority of students (75/124 = 60.48%) were concentrated in the mid-frequency range of 20-40, with 33 high-frequency positive interactions of >40.

As for the dimension of the degree of independent learning in the attribute stratum, the average length of students' weekly independent practice outside class was 181 minutes, with a median close to 184 minutes, i.e., half of the students practiced independently for more than 3 hours per week, and their practice habits were relatively stable, which can be found to show a very standard left-narrow-right-broad normal distribution in Fig. 6(d), and there were only 2 students who practiced <1h per week, and >4h for 19 students, the overall degree of independent learning of the students is good.

The peer evaluation is conducted on a 4-point scale, with an average score of 3.25. The overall result falls within the "good" range. In the peer evaluation process, a significant number of students, 55 and 47 respectively, received "excellent" and "good" ratings. This indicates that there is a favorable atmosphere for collaboration and mutual recognition among students.

In the pre- and post-course music skills test comparisons, there was an average gain of 2.42 points, with the majority of students receiving small to medium gains in performance from 0-4 points, and 13 students even improved by >4 points. However, there were individual students (12/124) who showed a small regression.

As for the expectation match of students in the goal expectation level, the average value is as high as 89.83%, which means that the vast majority of students believe that the course learning achieves their expected goals, and only 12.4% of the students have an expectation match of <80%. This reflects a good match between the current curriculum and students' needs.

3.1.2 Factors influencing learning outcomes

In order to explore the correlation between the variables, this study used Pearson correlation analysis and the correlation matrix is shown in Figure 7.

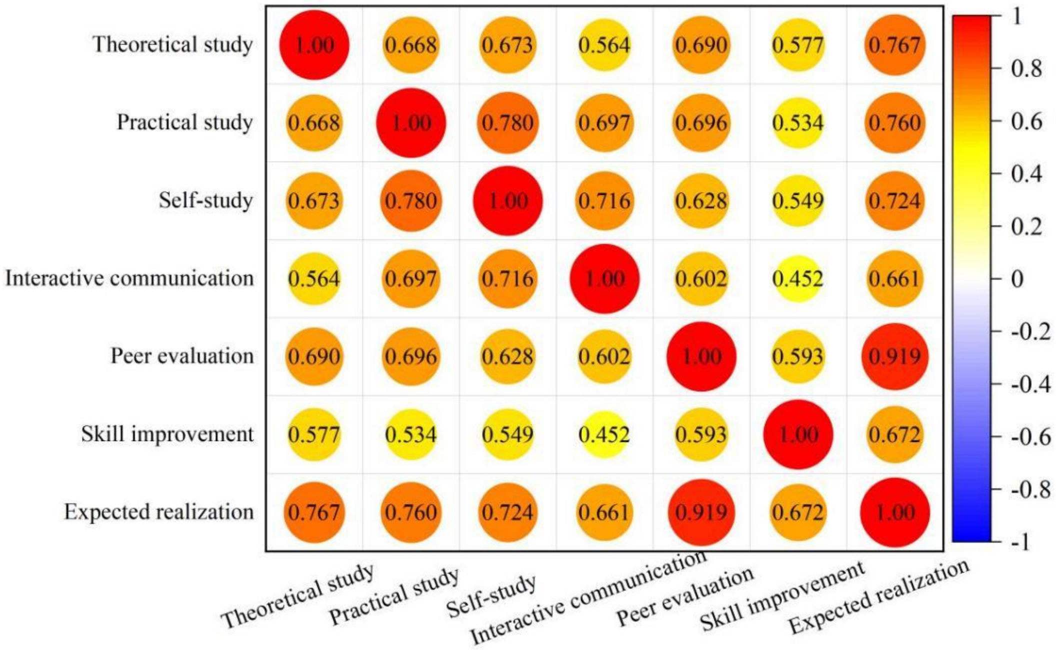


Figure 7: Correlation matrix of 7 music learning indicators

All coefficients in the table are positive and no negative correlation is seen. That is, there is a synergistic relationship between the seven oriented indicators of music learning under study. Among them, there is a high correlation between expectation fulfillment and peer evaluation, with Pearson's coefficient r=0.919. Meanwhile, there is also a stable positive correlation between theoretical and practical learning, with r=0.668, and the students who learn music solidly tend to be more expressive in singing and playing as well. And there was also a significant correlation between independent learning and interactive communication, r=0.716, indicating that students who were willing to spend time practicing independently were also usually more actively involved in rehearsal collaborations, which was similar in many dimensions.

3.2 Canopy-Kmeans clustering division based on students' music performance

3.2.1 3D Clustering Demonstration of Learning and Attribute Layers

Further, the clustering is mainly divided according to three indicators of the student learning layer, i.e., music theory final grade (theoretical learning), singing/playing score (practical learning), and frequency of cooperative rehearsal (interactive communication), which are X, Y, and Z axes, respectively, and the degree of self-directed learning in the attribute layer, i.e., the length of self-directed practicing hours outside the classroom, is introduced as a bubble size mapping. The obtained 3D clustering display based on learning and attribute layers is shown in Figure 8.

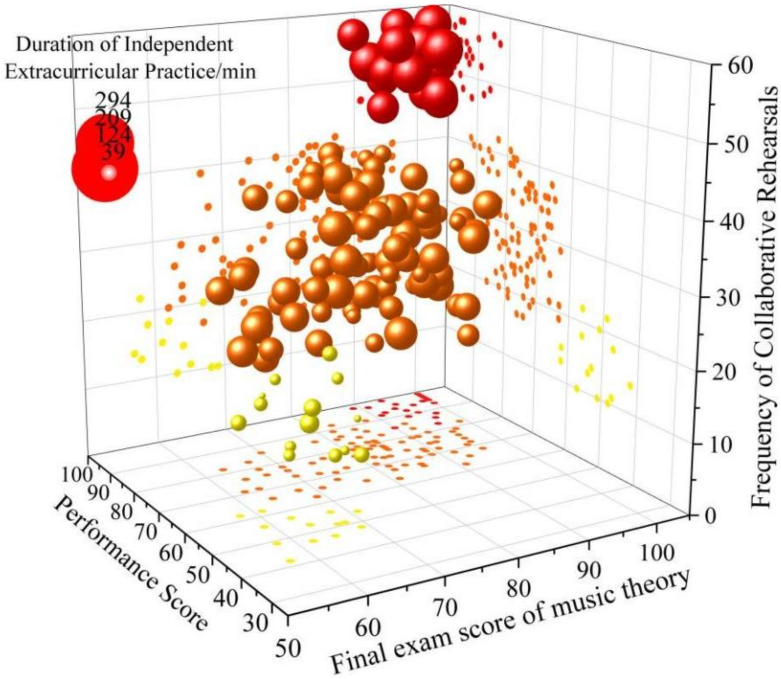


Figure 8: 3D clustering display of the learning layer and attribute layer

Three groups with significant differences were obtained based on Canopy-Kmeans clustering. Cluster 1 students (N=25) were well-rounded expressive learners, this group of students excelled in all indicators of the learning and attribute layers, with almost all of their theory grades and performance scores above 90, and the 25 students ranged from 45-60 on the Z-axis (frequency of collaborative rehearsals), which corresponded to the largest bubble of weekly independent practice hours outside of class of more than 4 hours per week. This group of students not only achieved both theoretical and practical excellence in their individual skills, but were also significantly higher than the overall level of collective cooperation and autonomous commitment.

Cluster 2 students (N=85) were the middle majority that made steady progress. This group was the largest, with all indicators in the middle range, with X-axis theory scores and Y-axis performance scores generally ranging from 60-90, with their cluster centers at 82.47 and 75.89, respectively, and with semester frequency of collaborative rehearsals centrally distributed around 18-42, and weekly autonomous practice hours corresponding to bubble sizes of between 125-209min. The students as a whole met the basic expectations of the course, with a solid grasp of theory, fair performance ability, and a normative level of frequency of participation and interaction.

Cluster 3 students (N=14) belonged to the hangman type in dire need of support, a small group of students who were significantly low on all indicators, corresponding to the small yellow bubble data point located in the lower left corner of the 3D graph, with average theory and singing/playing scores almost always below the 60 passing line, with an average semester frequency of collaborative rehearsals of only 15, and even less than 1.5 hours of weekly self-directed practice time.

3.2.2 Validation of Canopy-Kmeans Clustering Methods

In order to further verify the clustering classification accuracy obtained by the Canopy-Kmeans clustering algorithm in this paper, PCA combined with the K-means algorithm of PCA-Kmeans and the traditional Euclidean-K-means algorithm based on Euclidean distance that randomly selects the initial clustering centers are used as a comparative experiment for clustering the data respectively.

The data of 7 music learning indicators of students were fused to obtain a visual display of 124 students under 3 clustering algorithms, and the distance from data points to the clustering center was calculated as shown in Figure 9. Where the X-axis represents the clustering method, the Y-axis represents the clusters, and the Z-axis represents the distance from the data points to the center of mass.

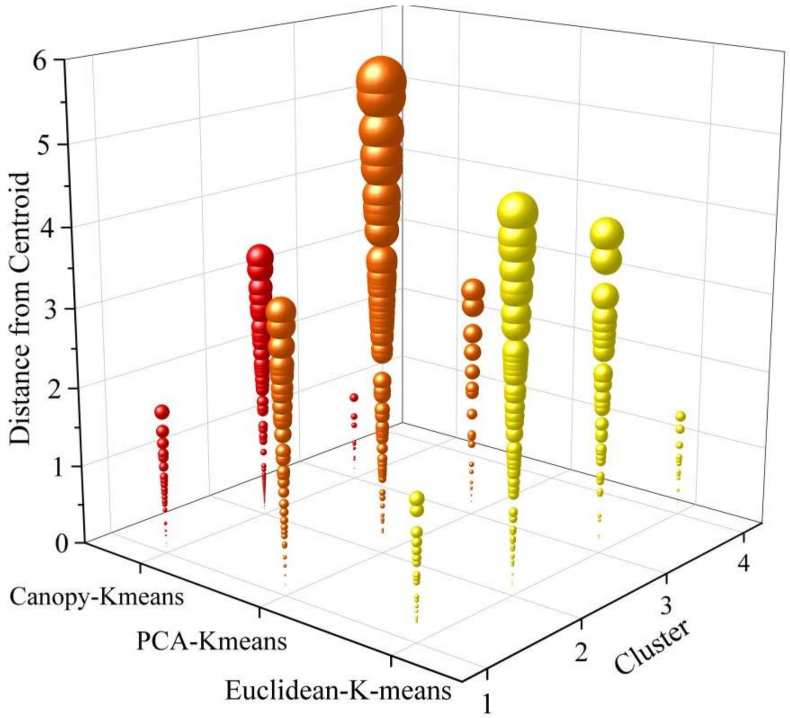


Figure 9: Display of clustering results under three algorithms

The distances from data points to their centers of mass within each category under the Canopy-Kmeans algorithm are all controlled at a low level: all the distances between cluster 1 and cluster 3 are almost within 1.5, and even in cluster 2, which has a wider range of range and a relatively more scattered distribution as in Fig. 8, the maximum distance of its data points is not more than 3.5. This indicates that the clusters have a compact structure and high cohesion, and the clustering results are clear and reliable. As confirmed by manual verification, the algorithm correctly classified 119 students with an overall accuracy of 95.97%. The algorithm in this paper adaptively determines the initial center of mass and the number of clusters through the Canopy coarse clustering stage, which effectively overcomes the shortcomings of the traditional K-means that is sensitive to the initial value and easy to fall into the local optimum. The robustness of the clustering process and the interpretability of the cluster structure are guaranteed on the multidimensional feature data of music learners.

In contrast, although the PCA-Kmeans method also obtained three clusters (N=33, 72, 19), the distance from some of its data points to the cluster centers was widely distributed, and the individual distances even exceeded 6.0, reflecting a looser clustering structure. The method finally correctly classified only 106 individuals with an accuracy of 85.48%. The four clusters generated by the Euclidean-Kmeans algorithm, N=20, 53, 36, 15, on the other hand, have more fuzzy boundaries and also have a large distance dispersion, with an accuracy rate of 87.10% (108 people correctly classified).

3.2.3 Performance analysis of clustered groups under multidimensional indicators

After verifying the superiority of the Canopy-Kmeans algorithm in terms of clustering consistency, accuracy and interpretability, the discovery of the students' data on the seven music learning indicators under the three clusters was continued, and their corresponding cluster class centers are shown in Table 1.

Table 1: 3 clusters center values corresponding of 7 music learning indicators

	Cluster 1(N=25)	Cluster 2(N=85)	Cluster 3(N=14)
Final grade of music theory	95.8	82.47	63.57
Performance score	93.24	75.89	52.5
Frequency of collaborative rehearsals	52.84	30.18	15
Weekly self-study practice duration/min	258.36	174.19	83.67
Peer evaluation	4	3.45	2.60
Comparison of music skills before and after the course	1.86	2.67	0.36
Expected matching degree/%	97.84	92.37	71..37

For Cluster 1 students (N=25) a well-rounded group of honors students. Their cluster centers for theory grades and singing/playing scores were 95.80 and 93.24, respectively, while maintaining a very high frequency of collaboration and extracurricular practice commitment, with cluster centers for both metrics of 52.84 and 258.36, respectively. Their peer ratings reached all 4's-“excellent”-and their average expected match was a high 97.84%. Although the average improvement of their skills is only +1.86 points, not even up to the average level of 2.42, it is because this kind of students start from a higher starting point, the space for progress is relatively convergent, and their learning has already entered the stage of stable refinement.

Cluster 2 students (N=85) represent the middle ground of steady growth. They were at an intermediate to advanced level in theory and practice, with cluster centers of 82.47 and 75.89, a total of 30 collaborative rehearsals a semester in frequency, and about 2.9 hours of practice per week. This group of students had the highest skill gains of the three categories, averaging 2.67, indicating that the course work provided them with significant ability growth. Peer evaluation (3.45) and expectation matching (92.37%) were also in the good range, showing an overall virtuous learning cycle of rewarding input, collaborative feedback, and basic achievement of goals.

Cluster 3 students (N=14), on the other hand, belonged to the supportive group that needed urgent attention. Their indicators were significantly low, with weak theoretical and practical performance averaging only 63.57 and 52.50, collaboration frequency of only 15 times a semester, and practice hours of approximately 1.4 hours per week, again significantly inadequate. Peer evaluations (2.6) were relatively tepid, with weak skill gains of 0.36 and even some negative regressions, and slow learning progress, with an expected match of 71.37%.

4 Role change of music teachers

Based on the three groups of students with different characteristics revealed by the cluster analysis in the previous section, the direction and path of the music teacher's role transformation will be discussed in the next section. Combined with the requirements of the new music curriculum standards, it will be elaborated how teachers can promote the all-around development of each learner through instructional design, practical guidance and evaluation feedback on the basis of respecting the differences of students.

According to the requirements of the new music curriculum standards, the role of college music teachers has gradually changed with the implementation of the new music curriculum standards, teachers in the teaching process need to be clear about their own professional positioning, the main body of teaching from the teacher to the students, fully respect the interests and needs of the students, through the classroom design, to guide the students to participate actively in the process, and the application of diversified teaching methods to enrich the process of knowledge transfer, to avoid the teacher's entire statement, the students are not the only ones who have been taught. Avoiding the teacher's entire statement, students passive listening to the whole process. For example, you can design the next class in advance of the relevant issues, give students the task of collecting relevant information, make full preparation before class, in the classroom, you can first let the students communicate with each other, and then the teacher to supplement, which can greatly improve the teaching effect. Students participate in the whole process, this sense of participation can greatly mobilize the students' enthusiasm and initiative, prompting changes in the traditional teaching methods and teaching concepts, students can be comprehensive development.

Nowadays, for music education in colleges and universities, the focus should be on helping students with musical creativity, to encourage students to put into practice as much as possible, in practice, constantly honing progress, and strengthen the music literacy is also an important part of the comprehensive quality. With the implementation of the concept of the new curriculum standards, the traditional single passive and old-fashioned teaching methods are no longer applicable, as a teacher in the context of today's times, should follow the changes of the times, keep up with the new concepts, new policies, abandon the past single passive teaching methods, to the new curriculum standards as the core theoretical guide, fully aware of the importance of the students' knowledge of learning combined with classroom practice, to change the traditional teaching in the The new curriculum standard is used as the core theoretical guide, fully aware of the importance of combining students' knowledge learning and classroom practice, changing the traditional teaching form of “teacher-led knowledge teaching”, and clarifying the role of the teacher as the designer of the music curriculum and the innovator of the teaching plan, constructing a new curriculum system for music education, and making behavioral changes to lay the foundation of promoting the multi-faceted development of students.

5 Conclusion

The study takes music learners' multidimensional data as an entry point, and reveals the characteristics of learner group differentiation in music education in the age of artificial intelligence through feature construction-clustering-population analysis, echoing the role transformation of music teachers. Based on the analysis of seven indicators of 124 students, the Canopy-Kmeans clustering method successfully classified them into three groups with an accuracy rate of over 95%. Among them, 25 "performance-oriented" students were almost uniformly ahead in all dimensions, and they excelled in the three theoretical indicators of learning level (95.8 points), practical indicators (93.24 points), and collaboration frequency (52.84 times). The 85 “growth” students comprise the majority of the students, and they showed the most significant and steady progress in terms of skill improvement (+2.67); The 14 “potential” students, on the other hand, were significantly lower on a number of indicators, especially in terms of the length of independent practice (83.67 minutes/week) and the frequency of collaboration (15 times), revealing both motivational and methodological challenges.

In this regard, the role of the music teacher must shift from uniform instruction to categorical support. Teachers should be guides to expressive students, providing them with more challenging tasks and creative space; teachers should be facilitators to help them break through the plateau period through structured feedback and goal dismantling; and teachers should be supporters to potential students, building up their practicing habits, enhancing their collaborative experiences, and gradually reshaping their learning confidence and rhythm. For potential students, teachers need to be proactive supporters to establish students' practicing habits and enhance their collaborative experience, so as to gradually reshape their learning confidence and rhythm.

About the Author

Dr. Pifen holds a Ph.D. in Musicology and is a high-level talent introduced by Zhaoqing University. She is currently a full-time faculty member at the School of Music and Dance at Kashgar University, where she is engaged in educational support work in Xinjiang. Additionally, she serves as the Deputy Director of the Greater Bay Area Aesthetic Education Association. Her primary research areas include vocal performance and teaching, AI technology empowerment in music education, music therapy, and multicultural integration. She has guided students to win provincial-level awards five times and has been honored with over ten titles, including “Outstanding Instructor.” Dr. Pifen also serves as a judge for international music competitions and has presided over and participated in multiple provincial and ministerial-level research projects. She has co-authored national planning textbooks and published several journal articles.

Wei Zheng is an Associate Professor at the School of Music, Linyi University. She holds a Ph.D. in Musicology and serves as a supervisor for master’s degree students. She is a member of the Asian Music Research Society and a committee member of the Technical Department of the Piano Society under the Linyi Musicians Association. Her main research areas include piano performance and pedagogy. She has presided over and participated in multiple provincial- and ministerial-level research projects, published three monographs in the field of piano pedagogy, and authored more than ten academic papers in domestic and international core journals.

References

[1] Fahimirad, M., & Kotamjani, S. S. (2018). A review on application of artificial intelligence in teaching and learning in educational contexts. International Journal of Learning and Development, 8(4), 106-118.

[2] Ghnemat, R., Shaout, A., & Abrar, M. (2022). Higher Education Transformation for Artificial Intelligence Revolution: Transformation Framework. Int. J. Emerg. Technol. Learn., 17(19), 224-241.

[3] AlAfnan, M. A., & Al, M. (2024). Taxonomy of educational objectives: Teaching, learning, and assessing in the information and artificial intelligence era. Journal of Curriculum and Teaching, 13(4), 173-191.

[4] Al-Zyoud, H. M. M. (2020). The role of artificial intelligence in teacher professional development. Universal Journal of Educational Research, 8(11B), 6263-6272.

[5] Tomaskinova, J., & Tomaskin, J. (2024). Unlocking the future of education: Empowering educators with AI by overcoming professional development challenges. In ICERI2024

Proceedings (pp. 10633-10642). IATED.

[6] Zhou, S. (2020). Analysis of the professional development of teachers in the era of artificial intelligence. Advances in Education, 10(3), 339-345.

[7] Ismail, A., Aliu, A., Ibrahim, M., & Sulaiman, A. (2024). Preparing teachers of the future in the era of artificial intelligence. Journal of Artificial Intelligence, Machine Learning and Neural Network, 44, 31-41.

[8] Bonales-Daimie, G., Martínez-Estrella, E., & Sierra-Sánchez, J. (2025). Evolution of the Teaching Profile and the Emergence of New Professional Roles in the Age of Artificial Intelligence (AI). A Perspective from Teachers, Students and Professionals. Pixel-Bit, Revista de Medios y Educacion, (73).

[9] Miao, Y., & Yao, Y. (2020, March). Professional development of college teachers in the era of artificial intelligence: role rebuilding and development path. In International Conference on Application of Intelligent Systems in Multi-modal Information Analytics (pp. 618-626). Cham: Springer International Publishing.

[10] Gao, R. (2025). The Erosion and Reconstruction of Teachers' Authority in the Context of Artificial Intelligence. Journal of Theory and Practice in Education and Innovation, 2(2), 26-33.

[11] Tao, B., Díaz, V., & Guerra, Y. (2019). Artificial intelligence and education, challenges and disadvantages for the teacher. Arctic Journal, 72(12), 30-50.

[12] Acil, T. (2024). Re-Thinking Boundaries: The Evolution and Impact of AI in Music and Soundscapes. AVANT. Pismo Awangardy Filozoficzno-Naukowej, (2), 1-15.

[13] Merchán Sánchez-Jara, J. F., González Gutiérrez, S., Cruz Rodríguez, J., & Syroyid Syroyid, B. (2024). Artificial intelligence-assisted music education: A critical synthesis of challenges and opportunities. Education Sciences, 14(11), 1171.

[14] Cheng, D., & Qu, X. (2025). Intelligent generation and optimization of resources in music teaching reform based on artificial intelligence and deep learning. Scientific Reports, 15(1), 30051.

[15] Zhao, X., Guo, Z., & Liu, S. (2021). Exploring key competencies and professional development of music teachers in primary schools in the era of artificial intelligence. Scientific Programming, 2021(1), 5097003.

[16] Luo, M., & Liang, S. (2025). Resistance or resilience? University music teachers' active learning intention in response to the emergence of artificial intelligence. Acta Psychologica, 261, 105775.

[17] Li, Y., Chai, Z., You, S., Ye, G., & Liu, Q. (2025). Student Portraits and Their Applications in Personalized Learning: Theoretical Foundations and Practical Exploration. *Frontiers of Digital Education*, *2*(2), 1-17.

[18] Alian, M., & Al-Naymat, G. (2022). Questions clustering using canopy-K-means and hierarchical-K-means clustering. *International Journal of Information Technology*, *14*(7),

3793-3802.

[19] Fan, D., Fan, L., Zhao, F., Sun, H., & Min, X. (2024, December). An Online Learner Clustering Algorithm Incorporating Autoencoder And Canopy-Kmeans. In *Proceedings of the 2024 2nd International Conference on Information Education and Artificial Intelligence* (pp. 627-633).