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

Analysis of the Teaching Model of Opera Singing Styles and Performance Skills Based on Data Mining Technology

Haotian Zheng[1,*]

[1] Conservatory of Music, Sahmyook University, Yichang, Hubei, 443000, China

SUMMARY: *In this research, basic music theories and notations are integrated to extract parameters of features for operatic singing concerning tonal dynamics, pitch, and breathing control. Pedagogic strategy is devised for students learning vocal performance in tertiary institutions, and a data mining service system architecture is designed to provide efficient informatization service delivery. Ninety vocal-performance students studying in the undergraduate program act as the subjects of the research and are divided into three teaching groups according to different pedagogic methods used for instruction. An analysis of the individual case study illustrates the implementation process of the proposed instruction method and measures the results obtained from the overall teaching practice. Compared with teaching group G1, there is a 30.00%, 36.66%, and 36.67% increase in pitch, rhythm, and soprano, respectively, for teaching group G3; compared with teaching group G2, there is a 43.33%, 33.33%, and 33.33% increase in pitch, rhythm, and soprano, respectively. The developed pedagogic method helps identify and adaptively adjust vocal problems in an accurate manner using the database to accumulate and optimize teaching resources continuously.*

KEYWORDS: *opera singing; feature parameter extraction; data mining; vocal performance skills*

1 Introduction

Opera is an important part of human music and art culture, which is a comprehensive art integrating music, drama, poetry, dance and stage art [1, 2]. Its teaching mode has always relied mainly on the teacher's experience, focusing on the teacher's performance, demonstration and modeling, but lack of accurate judgment and personalized guidance for the emotional treatment of singing style, performance skills, voice changes and other aspects. In order to improve personalized learning experiences and maximize students' overall performance capabilities, knowledge discovery approaches have become more common when considering studies on pedagogy and performance in opera singing

By making use of the analytics of computational methods, a large amount of information can be acquired based on learners' learning behaviors and performances. Teachers can make use of this kind of data to study learners' learning styles and their behavioral characteristics [3, 4]. For instance, through investigating learners' learning behaviors, the active periods of the learners' behaviors can be determined, thus allowing teachers to organize course materials and class arrangements in a way that is suitable for students' demands [5, 6]. Besides, computational analysis methods can help teachers find out what kind of learning and performance style learners prefer in opera singing, which will provide useful insights into

*17671265190@163.com
https://doi.org/10.65102/is2026185
Published: 30 April, 2026

organizing teaching materials [7, 8]. Through studying students' singing performances and evaluations in opera training, information like preferred styles can be learned about. In the teaching process, teachers can combine students' interests and preferences to choose appropriate opera materials and teaching methods to improve students' learning enthusiasm and participation [9, 10]. Data mining technology can also help teachers personalize the teaching content and provide targeted learning resources and exercises based on students' mastery of opera singing styles and performance skills to better meet students' learning needs [11-13].

The research, based on the fundamental concepts of music theory, examines feature criteria, including tone intensity, fundamental frequency, and breath, as well as comparing the comparative strength and weakness of students' singing voices. In order to solve the problem of inequality in music education resources, an analytic-based teaching support system is suggested. Furthermore, an analytic-based opera pedagogic framework is designed for the vocal-music students at the college level. By using two cases as examples, the process of applying the teaching models is illustrated. Through two cases, it can be seen that the use of this analytic technique will help students in better organizing their pitch, timbre, and dynamics. Finally, the performance of three different versions of the teaching models is examined.

2 Data mining-driven opera teaching model for college vocal performance majors

Amid the fast expansion of digital music techniques and artificial intelligence, the use of analytical mining methods in music education has become increasingly common, and its value in the instruction of university vocal performance students is especially prominent. As a highly integrated stage art, opera requires scientific approaches and systematic feedback to improve learners' command of singing style and performance techniques. Therefore, this study proposes an opera practice teaching framework for vocal performance students in higher education that incorporates analytical mining methods, establishes a scientific training and assessment mechanism, and advances the shift of opera instruction from conventional experience-based practice toward an evidence-oriented approach.

2.1 Fundamentals of music theory

2.1.1 Scientific Tone Notation and MIDI Pitch

The tones associated with different registers are generally represented through scientific pitch notation. In this notation system, fixed pitches are marked by the letters (C, D, E, F, G, A, B) together with an Arabic number. For instance, the internationally standardized pitch A4 has a fundamental frequency of 440 Hz, whereas C4 corresponds to 261.63 Hz.

Because auditory perception of pitch follows a logarithmic relationship with fundamental frequency, the human ear perceives the interval change from A3 (220 Hz) to A4 (440 Hz) as equivalent to that from A4 (440 Hz) to A5 (880 Hz). To align pitch variation with auditory perception in a more linear way, music theorists proposed expressing pitch by using the logarithm of the base frequency according to the number of scale transitions. For example, the widely adopted MIDI standard transforms the base frequency F0 into the MIDI pitch number p, and the exact transformation is given in Equation (1).

$$p = 69 + 12\log_2\left(\frac{F_0}{440}\right) \tag{1}$$

2.1.2 Short Score

Simplified notation is a familiar notation for the domestic public, which is simple, clear and easy to understand, and has made a significant contribution to the popularization and promotion of Chinese music.

The relative heights of the tones in the simple notation are indicated by seven Arabic numerals, except for the semitones between 3 (mi) and 4 (fa), 7 (si) and 1 (do), the others are all whole tones, and the correspondence between the tone names and the singing names is shown in Fig. 1 for example.

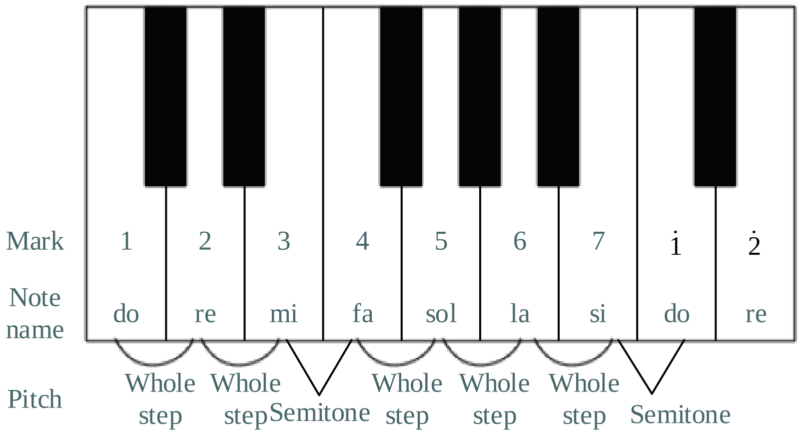

Figure 1: Example of the corresponding situation of phoneme and phoneme

Adding a black dot above the note means singing one octave higher, and similarly, adding two dots means singing two octaves higher; conversely, adding a dot below the note means singing one octave lower, and adding two dots means singing two octaves lower.

In addition to the seven Arabic numerals, the key signature is a very important symbol in the stave. The key sign is indicated by $1 = A$, $1 = F$, and so on. The stave can correspond 1 (do), 2 (re), 3 (mi), 4 (fa), 5 (sol), 6 (la), 7 (si) to the levels in the scientific key notation through the initial key notation.

The length of a note is indicated in simple notation by adding short horizontal lines after or under the note. An Arabic numeral indicates a quarter note, an Arabic numeral followed by a horizontal line indicates a half note, and an Arabic numeral followed by three horizontal lines indicates a whole note. An Arabic numeral followed by a horizontal line indicates an eighth note, an Arabic numeral followed by two horizontal lines indicates a sixteenth note, and an Arabic numeral followed by three horizontal lines indicates a thirty-second note.

The basic symbol for a rest is a 0. To indicate different lengths of rests, you can increase the number of zeros, add a short horizontal line under the zeros, or add a dot to the right of the zeros to mark them.

There is no concept of clef in the simple notation system, and pitch in simple notation is represented by notes and key signatures. The beat number in simple notation is the same as that in five-line notation, marked by a fraction. The key number and beat number of the piece are at the bottom left of the piece's name in the stave, and the author of the song is at the bottom right.

2.1.3 Short notation of songs

A complete score consists of the title of the piece, the key and beat numbers, the lyricist, and the note sequence. Note sequences contain a wealth of information about the song, not only retaining the melody, but also extracting the range of the highest and lowest notes, the range of the vocal range of the song, together with the key signature of the song. These note sequences may be divided into quarter notes, eighth notes, sixteenth notes, half notes, whole notes, among others. For example, a whole note is equal to two half notes, a half note is equal to two quarter notes, and a quarter note is equal to two eighth notes.

2.2 Feature Parameter Extraction

Sound signal is a time-varying signal, its waveform changes very quickly. However, if the time unit of observing the sound signal is narrowed down, it can be found that its waveform changes very slowly, which is called the short-time stability of audio. Using this characteristic, we can do the short-time distance processing of the sound signal, and extract the characteristic parameters of vocal singing such as the sound intensity, pitch and breath by the method of splitting frames.

2.2.1 Sound intensity

Tone intensity indicates the size of the sound and is determined by the amplitude of the sound signal. In vocal music teaching, it is not necessary to analyze each sample point, and the average amplitude in each frame of the signal can be obtained separately as the sound intensity parameter of this frame. In the system experiments of this paper, the set sampling rate of the sound signal is 44.1kHz, the number of sampling bits is 16bit, the frame length is 30ms, and the frame shift is 20ms. Each frame contains 1332 sampling points and 441 overlapping points. assume that each signal frame is denoted by $S_n(m)$, where $m = 0,1,\cdots,M-1$; $n = 0,1,\cdots,N-1$; N is the total number of frames, i.e., the length of the volume intensity curve; M is the frame size. The volume intensity curve is defined as:

$$Mag(n) = \frac{1}{M}\sum_{m=0}^{M-1}|S_n(m)|, n = 0,1,\cdots,N-1 \tag{2}$$

2.2.2 Pitch

Pitch indicates the height of a sound and depends on the frequency component of the acoustic wave, which can be estimated from the spectrum. Short-time frames can be extracted as described above for analyzing short-time energy, and the fundamental frequency can be determined by spectral analysis. Since the fundamental frequency is the most critical factor influencing pitch, the pitch of a performed vocal signal and a reference vocal sample may therefore be evaluated comparatively as long as the fundamental frequency is compared.

The detection of fundamental frequency is generally classified into three types: waveform estimation, correlation analysis, and transformation-based approaches. Within the vocal-assessment framework proposed in this study, the cepstrum approach from the transformation family is adopted to derive fundamental-frequency parameters. It is a conventional algorithm for detecting fundamental frequency that applies cepstral characteristics of the acoustic waveform to identify periodic excitation information associated with the vocal gate.

The sound $s(n)$ is obtained by excitation of the soundgate pulse $e(n)$ filtered by the channel response $v(n)$, i.e.

$$s(n) = e(n) * v(n) \tag{3}$$

Let the inverse spectra of $s(n), e(n), v(n)$ be $\hat{s}(n), \hat{e}(n), \hat{v}(n)$, respectively, and then there are

$$\hat{s}(n) = \hat{e}(n) + \hat{v}(n) \tag{4}$$

It can be seen that the fundamental information in the cepstrum domain is relatively separated from the channel information. A simple cepstrum filtering method can separate and recover $e(n)$ and $v(n)$, and the fundamental period can be found out based on the excitation $e(n)$ and its cepstrum characteristics.

The cepstrum analysis algorithm is more complicated, but its effect of base tone estimation is better. The cepstrum method of fundamental frequency detection is shown in Fig. 2, where the window function is chosen as Hamming window; $|S(w)|$ is the amplitude spectrum of the sound signal; $Ln|S(w)|$ is the logarithmic amplitude spectrum. After the Fourier inverse transform can get the signal cepstrum, from the cepstrum waveform can be estimated from the fundamental tone period.

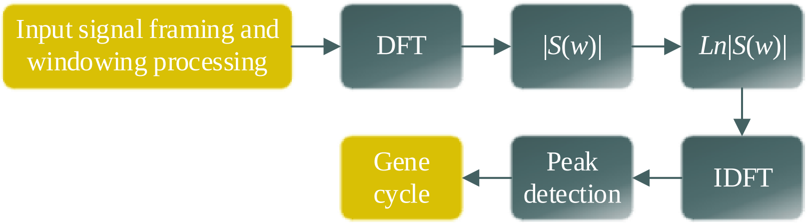

Figure 2: Fundamental frequency detection by cepstral method

2.2.3 The breath

Singing is a performing practice in which the human body serves as a sounding board. During vocalization, the human body and spirit should jointly achieve and sustain the condition needed for phonation. Proper breathing is the core force that supports the singing voice, so breath plays an essential role in singing. The measurement of respiratory control is a comparative evaluation procedure, and the smoothness of breath can be measured by calculating the standard deviation of an acoustic waveform. The standard deviation is a measure of how far a set of values is spread around the mean. A large standard deviation represents a large gap between most values and their mean, while a small standard deviation represents most values lying closer to the mean. Given a vector X, the standard deviation operator $std(X)$ can be written as follows:

$$std(X) = \left(\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right)^{\frac{1}{2}} \tag{5}$$

where n denotes the number of sampling points; $\bar{x}$ denotes the average amplitude.

2.3 Opera Practice Teaching for Vocal Performance Majors in Colleges and Universities

2.3.1 Train students' musical perception and construct musical polyphonic thinking

With regard to the practical application of the theoretical aspects of singing, along with the educational content used in the learning process in the vocal studies programs of higher education institutions, such information has evolved with respect to the Western music philosophy. Hence, while teaching the subject in practice, the tutors responsible for educating the specialized voice students in colleges and universities will be expected to employ opera-based methods to instill an understanding of musical awareness among the learners to enable them to enjoy the differences in Chinese vocal art and Western vocal art through practical participation in learning operas. The use of data mining methods will result in the development of the stylistic features databases that cover a wide range of musical performers and genres in order to classify the educational materials in music teaching. In today's art fusion background, the targeted fusion appreciation of different categories of works can assist in the overall enhancement of one's own musical perception ability. After all, for college vocal performance majors, starting from the cultivation of music perception ability as a foundation, can strengthen the overall practical experience of students, so that students slowly delve into the structure of the opera, and in the future to try complex, difficult performance works, to improve skills. For example, the teacher chooses the classic “Madame Butterfly” and “Sister Jiang”, which are two colorful Chinese and Western vocal opera works, and lets the students analyze the structural composition, tonal change, and timbre display of the two in the comparative learning mode, combining with the knowledge they have already learnt, so that the students can combine the teaching with the examples and form the musical polyphonic thinking.

2.3.2 Strengthening students' internal experience and improving their secondary creative ability

Opera itself belongs to the western music performance form, with strong practical, performance and comprehensive performance characteristics. Therefore, in the teaching of vocal performance majors in colleges and universities, teachers can start from the connotation experience of opera if they want to improve the existing teaching form and strengthen students' connotation knowledge of systematic vocal performance through opera practice teaching. The parameters of students' emotional expression in the process of interpretation are collected, and the excellent cases in the database are matched for similarity, and the students are guided to try to create secondary works on the basis of the deep understanding of the spirit of the original work, according to the results of the data analysis. This can not only exercise the students' performance skills and enhance their performance talents, but also understand the unique charm of opera consciousness from specific opera works and improve their secondary creation ability, so as to pave the way for the next professional learning and development. At the same time, the teacher organizes a comparative workshop on the interpretation of works, and with the help of visual charts of data mining, helps the students to recognize the artistic effects brought by different treatments. In the long run, students will naturally be able to know how to use different performance methods according to the emotions of the actual works, and start the actual opera practice to show their real performance level.

2.4 Data Mining Interactive System Architecture

To make better use of online educational resources, lower extra instructional costs, expand the value extraction of educational data, and apply the results more effectively to teaching and learning activities, this study designs an interactive architecture driven by data mining to support the classification and management of music curriculum resources. Within the overall framework of the data-mining-oriented information management platform, the front-end application is divided into two main modules, namely the learner side and the instructor side.

Their core purpose is to help users retrieve learning materials rapidly and conveniently through a search-engine mechanism developed on the basis of the open-source Lucene project. Meanwhile, search records can be stored efficiently. The recommendation component mainly relies on the database to capture user-related needs and then deliver suitable learning materials to users automatically. As a principal platform for informatized music education, the system also supports blog and forum functions, allowing users to share useful instructional content through the platform at any time and from any location. In this study, after establishing the overall framework, a more refined structural design is developed, as illustrated in Figure 3. At the infrastructure level, the entire platform is organized around the user interface to support interactive activities; in other words, the system is designed to enable users to complete a series of teaching-related operations, in which the user plays a central role. By relying on a real-time search engine, users are able to obtain comprehensive and updated music-learning materials. It should also be noted that the index table of the search engine remains in a continuously updated state and is stored in memory, so users can gain access more quickly and conveniently. The user-information database gathers data generated from interface operations, processes and categorizes these data efficiently, and then forwards them to the user model and recommendation module. The recommendation module subsequently pushes music-learning resources of potential interest to users through the platform, thereby ultimately achieving real-time interactive delivery of music teaching materials.

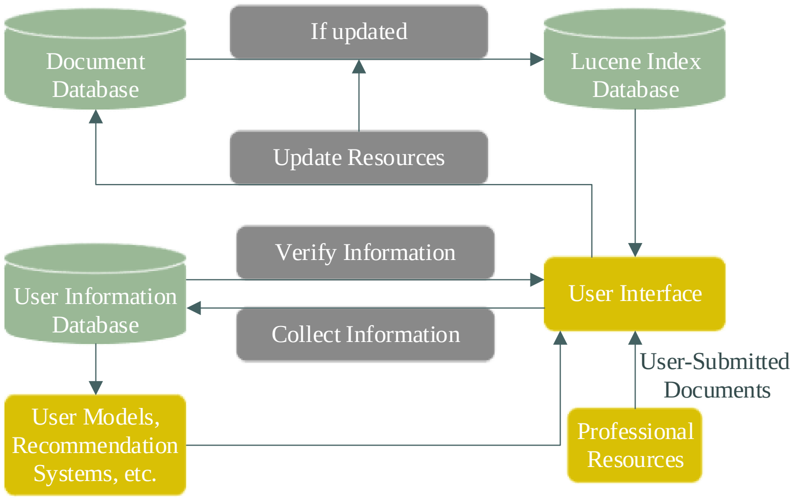

Figure 3: Architecture of data mining interactive information management system

3 Empirical analysis of the teaching mode of vocal performance majors in colleges and universities

In this study, 90 third-year undergraduates enrolled in vocal-performance programs at universities participated in the instructional experiment. Their mean age was 21, and the male-to-female proportion was 1:1. All participants had some prior background in vocal music. Together with the outcomes of the pretest on vocal ability, the students were randomly assigned to three groups, with 30 students in each group. Group G1 received a conventional instructional approach without the support of extra digital analytic tools, and the instructor delivered a total of 10 teaching hours. Group G2 employed the analytic method proposed in this study for independent learning and adjustment. Under instructor supervision, the teacher offered feedback only at key stages, while the students completed 10 credit hours of self-directed study. Group G3 was taught by applying the instructional framework developed in this study, including 6 credit hours of teacher-led instruction and 4 credit hours of student self-study, for a total of 10 credit hours.

3.1 Analysis and discussion of the results of voice acoustic testing

3.1.1 Tone map test results

One typical lyric tenor A and one typical lyric soprano B in group G3 were selected for individual analysis, and the results of their range comparison are shown in Figure 4. It can be seen that A's range map presents a narrow and long shape, with long horizontal line distances, a natural physiological range of nearly 50 semitones, a very wide range, and a convenient soprano with a strong vocal ability. While B's range map presents a wider and thicker type, with a natural physiological range of nearly 36 semitones, and compared to A's range map, the lateral distance is not long but the upper and lower line distances are farther apart, which indicates a greater variation in the dynamic range of the intensity of its sound.

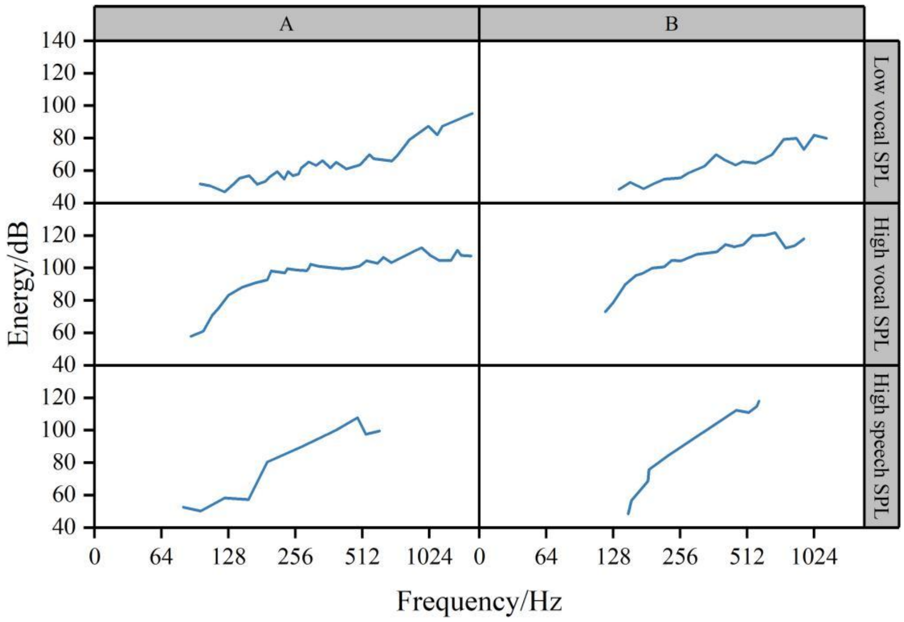

Figure 4: Results of vocal range comparison

Through the visualization of the pitch range map, students are able to visualize the distribution and boundaries of their pitches and identify their strengths and weaknesses. This clear perception of pitch ranges helps them to more accurately grasp their pitch choices in singing and enhance their musical perception.

3.1.2 Resonance peak results

The comparison of the sound spectra of subjects A and B is shown in Figure 5. It can be seen that A had a clear, focused “singer resonance peak” at 4000 Hz, while B had two “singer resonance peak” spikes at 3500 Hz and 4000 Hz, which did not merge into a peak as well as A. This may have some relationship with the fact that the cavity was not opened enough and the epiglottis was too small. This may be related to the fact that the cavity is not open enough and the epiglottis is too small.

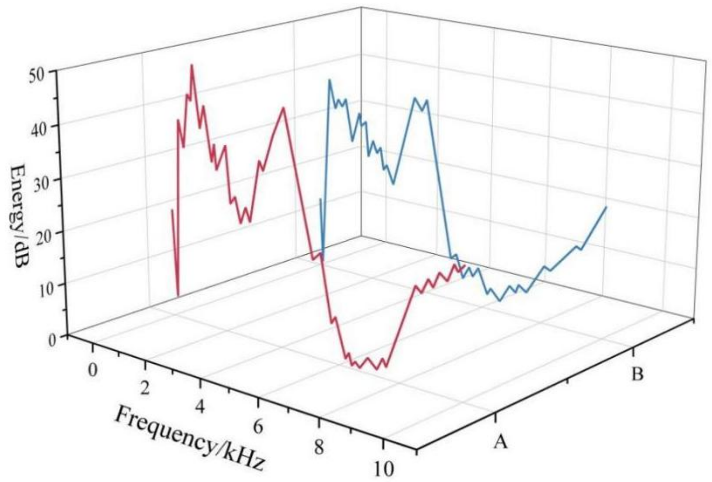

Figure 5: Acoustic spectrum comparison results

Within the instructional framework developed here, the incorporation of analytic tools can support students in perceiving timbre variation while simultaneously carrying out targeted cavity adjustment and resonance practice, thereby gradually enhancing the stability of high-frequency resonance and integration and further improving musical expressiveness.

3.1.3 Band energy analysis

Energy reflects sound intensity, as well as the degree of airflow produced by the singer and the magnitude of vibratory amplitude. In this study, three representative opera excerpts numbered O1-O3 were extracted using Praat software for acoustic analysis. The results of the analysis are presented below. The results of the section energy comparison between Subjects A and B when singing the three representative opera selections are shown in Figure 6. The average energy values of the three operatic excerpts of subject A ranged from 68.5 dB to 72.9 dB, and each of the three operatic excerpts had an energy concentration area, which was roughly 65 dB to 75 dB for O1, 65 dB to 75 dB for O2, and 70 dB to 80 dB for O3. The average energy values of the three long tunes of subject B ranged from 70.6dB-71.4dB.The three operatic excerpts each have three regions of energy concentration, with O1 having a concentration of energy roughly between 75dB-85dB.O2 has a very high concentration of energy compared to the other two, with a concentration of energy roughly between 55dB-85dB.O3 has a very pronounced energy ebb and flow, with a concentration of energy roughly between 55dB-70dB and 75dB-85dB.

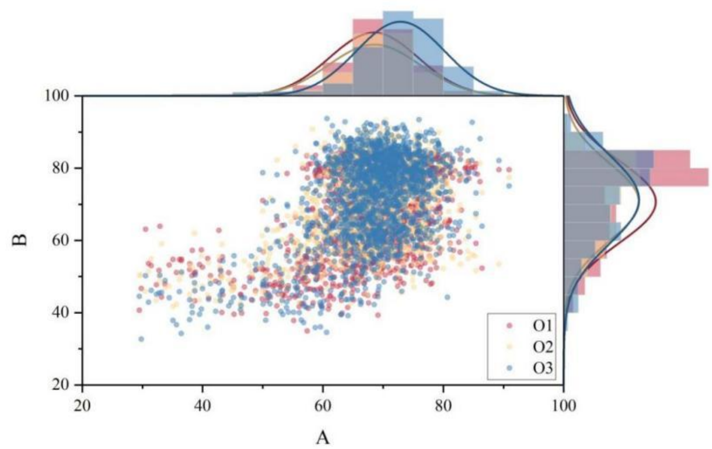

Figure 6: Results of energy comparison between bands

Through the visualization of energy curves, students are able to intuitively feel the correspondence between strength and breath usage. This sensitivity training to changes in strength helps them to be more comfortable in using the contrast between strength and weakness in their musical interpretation, enriching their musical expression.

In the above analysis process, the teaching resource database provides students with rich historical samples and parameter labels, enabling them to quickly identify voice types and match corresponding practice programs. At the same time, the longitudinal data accumulated in the database provides teachers and students with traceable progress records, supporting the continuous optimization of personalized feedback. It is the support of data mining technology that ensures that music characteristic parameter analysis can be maximized in teaching, so that the application of the tool and the cultivation of music perception ability form an organic combination.

3.2 Analysis of Teaching Effectiveness

Before the experiment, the participants' vocal performance was scored to identify their baseline singing level. Four evaluators assessed each student's singing from three dimensions: pitch, rhythm, and soprano. Each score was divided into five grades, with grade 1 representing the best performance and grade 5 the weakest. The mean score was then calculated as the participant's final result. After the experiment, the improvement in pitch, rhythm, and high-note performance before and after the intervention was statistically examined. "No improvement" indicated that the pretest and posttest scores remained at the same grade. "Improved by one level" indicated that the posttest score was one grade higher than the pretest score. "Improved by two levels" indicated that the posttest score was two grades higher than the pretest score. The improvement effect of different instructional modes was evaluated by examining the proportion of students whose performance increased under each teaching approach in a given dimension. The comparative results for improvement in pitch, rhythm, and treble performance across the three instructional groups before and after the experiment are shown in Table 1. According to the statistical findings, Group G3 achieved the highest improvement rates in pitch, rhythm, and treble, reaching 93.33%, 96.66%, and 90%, respectively. At the same time, only Group G3 included students whose performance improved by two grade levels, demonstrating that the instructional approach proposed in this study can produce highly favorable results. Groups G1 and G2 also showed similar improvement rates, with 63.33% and 50.00% in pitch, 60.00% and 63.33% in rhythm, and 53.33% and 56.67% in treble, respectively. In the pitch dimension, the percentage of improvement in the teaching group G2 with the intervention of the analyzing tool was lower than that in the teaching group G1, which, combined with the comparison with the dimensions of the teaching group G3, proves the importance of the teacher's timely intervention. The analytic tool alone does not guarantee the improvement of students' performance skills, and the combination of the teacher's on-the-spot explanation and targeted demonstration can bring the value of the tool into full play.

Table 1: Comparison results of the degree of improvement of the subjects

Dimension	Group	N	No upgrade		Upgrade 1 level		Upgrade 2 levels	
			N	Proportion	N	Proportion	N	Proportion
Intonation	G1	30	11	36.67%	19	63.33%	0	0.00%
	G2	30	15	50.00%	15	50.00%	0	0.00%
	G3	30	2	6.67%	21	70.00%	7	23.33%
Rhythm	G1	30	12	40.00%	18	60.00%	0	0.00%
	G2	30	11	36.67%	19	63.33%	0	0.00%
	G3	30	1	3.33%	25	83.33%	4	13.33%
Treble	G1	30	14	46.67%	16	53.33%	0	0.00%
	G2	30	13	43.33%	17	56.67%	0	0.00%
	G3	30	3	10.00%	24	80.00%	3	10.00%

4 Conclusion

This study makes effective use of database-based information methods to investigate in depth a music-curriculum framework supported by analytical mining techniques. Through instructional implementation, it demonstrates the full procedure of integrating analytic mining approaches into opera teaching.

(1) The examination of representative cases indicates that the analytic tools developed in this study can enhance students' vocal abilities and strengthen the growth of musical perception, while the teaching-resource database plays a foundational role in supporting the entire process of detection, analysis, and feedback.

(2) The overall teaching effect shows that the proportion of students in teaching group G3 who improved one grade in the dimensions of pitch, rhythm, and treble reached 70.00%, 83.33%, and 80.00%, respectively, and the proportion of students who improved two grades reached 23.33%, 13.33%, and 10.00%, respectively, which is the best performance among the three teaching groups. In the data mining technology-assisted teacher teaching model, the teacher's on-site guidance is a clear advantage.

About the Author

Haotian Zheng was born in Yichang, Hubei Province in 2000. I graduated from Wuhan Conservatory of Music with a bachelor's degree, majoring in Music Education (Vocal Education). I am currently a doctoral student at Samyuk University in South Korea, majoring in Vocal Opera.

References

[1] Goehr, L. (2014). The concept of opera. Chapter, 5, 92-133.

[2] Meyer-Dinkgräfe, D. (2014). Opera Criticism: State of the Arts and Beyond. Opera in the Media Age: Essays on Art, Technology and Popular Culture, 43-64.

[3] Na, W. (2020). A Data Mining Method for Students' Behavior Understanding. International Journal of Emerging Technologies in Learning, 15(6).

[4] Jie, W., Hai-yan, L., Biao, C., & Yuan, Z. (2017, June). Application of educational data mining on analysis of students' online learning behavior. In 2017 2nd International Conference on Image, Vision and Computing (ICIVC) (pp. 1011-1015). IEEE.

[5] Sun, X. (2024). Educational Technology Based on Data Mining: Mining Student Behavior Patterns to Optimize Teaching Strategies. International Journal of High Speed Electronics and Systems, 2540101.

[6] Li, S. (2024, September). Research and Development of a Student Behavior Analysis System Based on Big Data Mining. In 2024 3rd International Conference on Artificial Intelligence and Computer Information Technology (AICIT) (pp. 1-7). IEEE.

[7] Wang, T., Xiao, B., & Ma, W. (2022). Student behavior data analysis based on association rule mining. International Journal of Computational Intelligence Systems, 15(1), 32.

[8] Chen, L., Wang, L., & Zhou, Y. (2022). Research on data mining combination model analysis and performance prediction based on students' behavior characteristics. Mathematical Problems in Engineering, 2022(1), 7403037.

[9] Algarni, A. (2016). Data mining in education. International Journal of Advanced Computer Science and Applications, 7(6), 456-461.

[10] Zain, J. M., & Herawan, T. (2014). Data Mining for Education Decision Support: A Review. International Journal of Emerging Technologies in Learning, 9(6).

[11] Ojha, S. K., & Schan, K. K. (2024). The Impact of Data Mining on Curriculum Development: Tailoring Education to Student Needs. International Journal of Sciences and Innovation Engineering, 1(2), 37-42.

[12] Magdin, M., & Turcáni, M. (2015). Personalization of Student in Course Management Systems on the Basis Using Method of Data Mining. Turkish Online Journal of Educational Technology-TOJET, 14(1), 58-67.

[13] Papadogiannis, I., Wallace, M., & Karountzou, G. (2024). Educational data mining: A foundational overview. Encyclopedia, 4(4), 1644-1664.