Hindawi
Mobile Information Systems
Volume 2022, Article ID 2470777, 7 pages
https://doi.org/10.1155/2022/2470777

Research Article

A Data Mining-Based Evaluation Technique for Music Teaching

Zeng Qin [1,2]

[1] School of Music and Dance, Chengdu University, Chengdu 610106, China
[2] College of Education, Central Philippine University, Iloilo, Philippines

Correspondence should be addressed to Zeng Qin; zengqin@cdu.edu.cn

Received 21 February 2022; Revised 9 March 2022; Accepted 16 March 2022; Published 19 April 2022

Academic Editor: Mian Ahmad Jan

Musical data mining covers a number of methodologies to successfully apply data mining techniques for music processing, drawing together a multidisciplinary array of top experts. The field of music data acquisition has grown through time to solve the difficulties of obtaining and engaging with enormous amounts of music and associated data, such as styles, artists, lyrics, and reviews. In order to improve the quality of music teaching, a music teaching evaluation based on data mining is proposed. Data mining is becoming more widely accepted as a viable form of inquiry for analyzing data obtained in natural settings. More and more attention is paid to music teaching. Actual data is frequently inadequate, unreliable, and/or lacking in specific behaviors or patterns, as well as including numerous inaccuracies. Preprocessing data is a tried-and-true means of resolving such problems. Music teaching data is divided into three steps after preprocessing, that is, "object and object type," "music teaching data normalization," and "data integration." A model is built with a high-dimensional characteristic distribution and essential parameters of convergent teaching capacity. The experimental results show that the data mining method can be used for music teaching evaluation and has the advantages of short evaluation time, high accuracy, and clear indicators.

1. Introduction

Music education is an integral part of school art education. It is an important way to cultivate students' artistic accomplishment and aesthetic taste. As a special humanities subject, music teaching evaluation has many conflicts of value, such as the conflict in "evaluation index," conflict in "rating function," and conflict in "evaluation pursuit." The evaluation index includes "technical" and "emotional" conflict, the rating function includes "instrumental" and "developmental" conflict, and the evaluation pursuit includes "long-term benefit" and "short-term benefit" conflicts. "Music teaching evaluation" is an important stage in the process of music teaching in teachers and music learning in students. In a broad sense, music teaching evaluation covers all fields of music teaching. The function of teaching assessment is highly significant in the teaching course, especially for art teaching forms such as music instruction, which influences the teaching quality directly or indirectly. In view of the value and difficulty of music teaching evaluation, it is imperative to study the method of music teaching evaluation.

Some of the previous studies on music evaluation have been gathered. A lot of different algorithms have been introduced in the preceding research. Guo et al. [1] introduced the algorithm of a comprehensive evaluation of average value and AHP fuzzy comprehensive evaluation and realize it in Python. The results of the study suggest that the AHP fuzzy comprehensive evaluation algorithm is appropriate for teaching management and can increase the accuracy and efficiency of teaching quality assessment. In Wang et al.'s [2] study, in view of the problem of how to better improve the learning effect of computer courses for non-commissioned officers and cadets, practically and effectively evaluate the learning results, stimulate the learning motivation, improve the quality of classes, and finally achieve the improvement of their basic accomplishment in military informatization and in combination with their own teaching practice, the scoring model of courses has been reformed and substantial progress has been made. In this paper, a large number of data are collected and analyzed. The reform of the grading model and its influence on the quality of music teaching are further analyzed. In Xia et al.'s [3] study, a new music teaching evaluation model based on the weighted Naive Bayes algorithm is proposed. In addition, to improve the efficiency of the music instruction evaluation system, the weighted Bayesian classification incremental learning method is applied. In Espigares-Pinazo et al.'s [4] study, through the remote information processing platform, the use of automated data analysis programs in the teaching process was introduced. The application of the music ability evaluation and analysis program was proposed. This research paper was based on the data collected in the examination at the end of the course. The acquisition level of key musical ability was evaluated objectively using "K" which means classification technology.

Data mining, a new data analysis tool, has been widely used in many fields and achieved remarkable results. Data mining in music teaching outlines a number of methods for successfully applying data mining techniques to music processing. Its strong ability in data integration and analysis has become a hot research topic. The field of music information retrieval has grown through time to solve the difficulties of efficiently accessing and engaging with enormous collections of music and associated data such as genres, artists, lyrics, and evaluations. The music data is preprocessed before evaluation. The preprocessing of data includes different stages to evaluate the data. An experiment has been performed on 150 songs that have been selected randomly. The experimental data were compared with the weighted Naïve Bayes algorithm [3] and automated data analysis programs [4]. The results of the experiments suggest that the data mining method may be used to evaluate music training and that it has the advantages of a quick evaluation time, high accuracy, and unambiguous indicators. Therefore, this paper presents a music teaching evaluation method based on data mining.

The rest of the paper is as follows. After Introduction, the preprocessing method of music teaching data is discussed in Section 2, and then, the music data evaluation using data mining is discussed in Section 3. Next, the experimental verification of the data analyzed in the entire paper is presented in Section 4, and lastly, the conclusion of the paper is discussed in Section 5.

2. Preprocessing of Music Teaching Data

In the process of evaluating music teaching, data is retrieved from relevant databases or data files and then preprocessed to organize and modify part of the data that has to be processed. Finally, the processed data are integrated [5, 6] and then stored in the corresponding data warehouse. The data warehouse stores the data according to a certain organizational structure. According to the teaching data, it builds up to two independent platforms for teachers and students. After that, it mines the data effectively, finally analyzes the mining results, and forms an evaluation analysis report [7, 8]. To conclude, there are three steps to data mining in teaching evaluation: data preparation and preprocessing, data mining, and outcome analysis. A data mining technique is used in the assessment process to process and transform large volumes of teaching data, such as association rules, in order to better analyze this data. Therefore, the preprocessing of music teaching data mainly includes three steps: determining the mining object and the target category, normalizing the music teaching data, and preprocessing the discrete data.

2.1. Identify Mining Objects and Target Categories. In identifying the mining objects, it is an important step to recognize the goal of data mining and clearly define the problem. The final results of mining are unpredictable, but exploring the problem should be foreseen. Therefore, the fuzzy theory [9, 10] is used to identify the music teaching objects and target categories to be mined in order to provide a basis for the accurate processing of data.

Set $U_M(X) = \Lambda u_{a_i}(x_i)$ as the goal of music teaching, and set X as the value of membership function for interval set M. Each element x_i has only one fuzzy interval a_i corresponding to it in the fuzzy interval set M. $u_{a_i}(x_i)$ is the degree of the element x_i belonging to the fuzzy interval a_i.

Based on the above analysis, the fuzzy relational support of the music teaching dataset is obtained as follows:

$$W_i = \frac{\sum_{i=1}^{m} E_i}{U_M(X) + U_N(Y)}. \tag{1}$$

In formula (1), m is the number of music teaching objectives, E_i is the data value, and $U_N(Y) = \Lambda u_{nj}(y_j)$ is the membership function value of set Y to fuzzy interval set N. The music teaching indexes are as follows:

$$C_{ij} = \frac{U_M(X) + U_N(Y)}{2}. \tag{2}$$

Calculate the support degree and confidence degree of fuzzy association among data classifications using formulas (1) and (2); then, update the data classifications based on the calculation findings to establish a foundation for data mining.

2.2. Normalization of Music Teaching Data. According to the data classification adjustment in Section 2.1, all the data related to music teaching are tested, and the data related to each teaching are recorded in chronological order. However, due to data diversification, if not normalized, the effect of music instruction data mining will be harmed, and mining efficiency will be reduced. Therefore, this paper transforms the data of different teachers and different repertoire courses. The conversion formula of different teaching repertoire is

$$P_i = g_{ia} + g_{ib} + g_{ic}. \tag{3}$$

g_{ia} stands for teaching time, g_{ib} stands for number of students, and g_{ic} stands for teaching information. Calculate the teaching quality of each track course based on the converted content:

$$Z_k = \sum_{k=1}^{q} T_{ka} + \sum_{k=1}^{q} T_{kb} + \sum_{k=1}^{q} T_{kc}. \tag{4}$$

Among them, T_{ka} represents the number of samples, T_{kb} represents the proportion of teaching repetition, T_{kc} represents the students' receptivity, q represents the total number of courses, and k represents the random error of different teacher samples.

Taking track data as an example, the data after data cleaning is represented in the form of a set: $V = \{v_1, v_2, \cdots, v_s\}$, where $s = 1, 2, \cdots, S$, maps the set to a sample space which is composed of S sample points, where the sample mean can be calculated by the following formula:

$$E_v = \frac{Q_i \times Q^2}{\sum_{s=1}^{q} V_s \times Z_k}, \tag{5}$$

where Q_i represents the sample size and Q^2 represents the sample variance. When the sample size is high enough, the sample mean and variance tend to match the population expectation and variance, respectively.

Create a set of samples to be predicted $F = \{f_1, f_2, \cdots, f_w\}$, in which $w = 1, 2, \cdots, L$, the set of all the samples as a prediction sample. We should quantify the specific features of the samples in order to create reasonable sample predictions. The properties that have been quantified are referred to as sample indicators. If there are j indicators, the samples can be described by d vector, that is,

$$R = \{r_{1d}, r_{2d}, \cdots, r_{wd}\}. \tag{6}$$

Since, in practice, the sample data collected are often not those in [0, 1] closed intervals, these raw data should be standardized and averaged:

$$\bar{R} = \cdot \frac{\tau_{ij}(t)}{L}. \tag{7}$$

Of these, $\tau_{ij}(t)$ represents the data for item i of the L sample and j represents the average number of teaching tracks. Quantification of teaching characteristics is as follows:

$$R_s = 1 - \frac{\sum_{i=1}^{q} P_i}{L - 1}. \tag{8}$$

Extreme value standardization is calculated using the following formula:

$$C_s = \frac{\sum_{s=1}^{q} R_s \times d_f}{U}. \tag{9}$$

Among them, R_s denotes the teaching grade of the track, U denotes the number of main teaching tracks, and d_f denotes the value of students' comprehension ability.

2.3. Discretization Pretreatment. Integrate the music teaching data and use the relevant elements such as curriculum and student factors to establish the music teaching dataset and verify. The pretreatment process is as follows:

(i) Selecting music teaching samples

(ii) Estimating the number of teaching courses on the same day

(iii) Establishing training set and verification set by using influence factor and static factor

The process is as follows:

(i) Calculating the cumulative teaching time and teaching repertoire information of the course: curriculum teaching time and teaching content are set as influencing factors to analyze their relationship with the quality of music teaching. The formula for calculating the quality of music teaching shall be

$$T_i = t_1\chi + t_2\chi + \cdots + t_y\chi \tag{10}$$

In the formula, t_1 is the accumulative teaching time of the first day, y is he number of teaching tracks, t_y is the accumulative teaching time of the first track of y, and χ is the key symbol of the track.

Calculate the Pearson correlation coefficients between the quality of instruction and the parameters:

$$r_{x_i h} = \frac{S_{x_i h}^2}{S_{x_i x_i} S_{HH}}. \tag{11}$$

In the formula, h is the quality of music teaching, x_i is the influence factor, $S_{x_i h}^2$ and $S_{x_i x_i}$ are the standard deviation under the influence of covariance between h and x_i, and S_{HH} is the standard deviation of teaching quality evaluation. Based on the Pearson correlation coefficient, the y and χ with the maximum absolute value are selected as the final correlation coefficients to calculate the cumulative teaching program and teaching time.

(ii) Discretization of impact factors: the influence factors of music teaching evaluation are mostly continuous variables. Discretization of the influence factors can improve the ability of anti-interference, the completeness of information collection, and the effectiveness of teaching evaluation

(iii) Data conversion: the integration of different data information collected by teachers into an analytical data model is prepared for algorithms, which may require different analytical data models

(iv) Data algorithm mining: the aim of using appropriate mining algorithm is to build a data mining model. Firstly, we must choose the appropriate mining algorithm and use the appropriate program design software to implement the algorithm

(v) Analysis of the results of the excavated rules: explain and evaluate the mining results

3. Music Teaching Evaluation Based on Data Mining

Based on the above preprocessing of music teaching information, the index set G of music teaching evaluation includes 9 aspects: teaching attitude G_1, teaching content G_2, teaching art G_3, classroom structure G_4, classroom management G_5, teaching effect G_6, evaluation type G_7, student role G_8, and evaluation credibility G_9, which is recorded as $G = \{G_1, G_2, G_3, G_4, G_5, G_6, G_7, G_8, G_9\}$.

3.1. Music Teaching Data Clustering. Cluster analysis is a common method to study the classification of things. It classifies the similar things into the same class and the different things into different classes. What distinguishes it from classification is that it deals with things without having defined the categories in advance, even when there are several classes in all. Therefore, cluster analysis has no training samples as a basis for classification, and there is no theoretical classification principle; therefore, it can only be classified based on the nature of things' similarity [11, 12]. As a result, cluster analysis is based on the nature of items and the degree of approximation used to classify them. The clustering statistics are a measure of how near items are.

The statistical quantity of index cluster is similarity coefficient; the bigger the similarity coefficient is, the more similar the two indexes are. Therefore, the cluster principle states that indexes with a high similarity coefficient belong in the same class, while indexes with a low similarity coefficient belong in a different class [13]. The correlation coefficient is used to indicate the similarity coefficient of quantitative indexes.

Given that there are n variables to be clustered, which are recorded as $X_1, X_2, \cdots X_n$, and the existing z samples, which are recorded as $A_p, p = 1, 2, \cdots, z$, the correlation coefficient of X_a and X_b is calculated as

$$R_{ab} = \frac{\operatorname{cov}(X_a, X_b)}{\sqrt{DX_a}\sqrt{DX_b}}. \tag{12}$$

$\operatorname{cov}(X_a, X_b)$was the covariance of X_a and X_b, and DX_a and DX_b were the variance of X_a and X_b, respectively.

The similarity coefficients of X_a and X_b are expressed by c. When R is positive, then $c = R$; if R is negative, then the similarity coefficient is defined in the following two ways: $c = |R|$ or $c = R + 1$, generally using the former.

The statistical quantity of sample clustering is called distance. The shorter the distance between two samples is, the closer they are. Therefore, the basic premise of sample clustering is to group samples that are close together into one class and samples that are far apart into multiple groups [14, 15].

The Euclidean distance between sample A_1 and A_z is commonly used for distance indices.

$$d_{pz} = \sum_{\chi=1}^{n} \left(X_{p\chi} - X_{z\chi}\right)^2. \tag{13}$$

If the average distance is considered too important for values with large absolute values, the absolute value can be used as an example. The absolute value distance between sample A_1 and A_z is

$$\widehat{d}_{pz} = \sum_{\chi=1}^{n} \left|X_{p\chi} - X_{z\chi}\right|. \tag{14}$$

When the index units of the sample are different or the order of magnitude is different, the index shall be standardized before calculating the distance.

3.2. Music Teaching Evaluation Methods. In order to evaluate the convergence index parameters of music teaching evaluation, it is required to construct a model of high-dimensional characteristic distribution and key parameters of convergent teaching ability, such as the level of teachers and the degree of policy relevance. The data flow of convergence parameters of music teaching evaluation is given by constructing differential equation.

$$H_n = h\left(\widehat{d}_{pz} \cdot c\right) + \omega_n. \tag{15}$$

Among them, $h(\cdot)$ is the multivariate value function of teaching evaluation, and ω_n is the detection value of evaluation deviation.

The solution vector of music instruction assessment is calculated via correlation fusion in the high-dimensional feature distribution space, and the feature training subset δ_r of teaching ability evaluation is obtained, and the following constraints are met:

$$K_r = \operatorname{diag}(\delta_1, \delta_2, \cdots, \delta_r), \quad \delta_r = \sqrt{\lambda_r}. \tag{16}$$

Set λ_r as the conjugate solution of the statistical data of music teaching evaluation, which accords with the original value characteristic decomposition condition $B = \{b(\chi) \mid b(\chi) \in X, \|b\| \le d, t \in K\}$. The data flow calculation formula for music teaching assessment is generated based on the prior statistical test value and the statistical characteristic distribution sequence of a set of multivariate variables in music teaching evaluation.

$$\begin{aligned} c_{1x}(\tau) &= E\{x(n)\} = 0, \\ c_{2x}(\tau) &= E\{x(n)x(n+\tau)\} = r(\tau), \\ c_{kx}(\tau_1, \tau_2, \cdots, \tau_{k-1}) &\equiv 0, \quad k \ge 3. \end{aligned} \tag{17}$$

If $q = 200$, the level of teachers and the level of distribution of teaching resources of music teaching evaluation meet the $(200 + 1)$ maintenance functional conditions; that is, the teaching ability evaluation has a constrained solution, and the convergence conditions are

$$\Psi_x(\omega_n) = \ln \Phi_x(\omega_n) = \frac{{\omega_n}^2 \lambda_r}{2}. \tag{18}$$

Quantitative recursion of big data of music teaching evaluation is used to establish the control objective function of teaching ability prediction:

$$\max_{x_a,b,d,p} \quad \sum_{a\in A}\sum_{b\in B}\sum_{d\in D}\sum_{p\in P} x_{a,b,d,p} V_p$$
$$\text{s.t.} \quad \sum_{a\in A}\sum_{d\in D}\sum_{p\in P} x_{a,b,d,p} R_p^{bw} \le K_b^{bw}(S), \quad b \in B. \tag{19}$$

If the historical data of teaching ability distribution of quantitative recursive results of music teaching level is J_k, then the probability density function of the predictive judgment of music teaching is under the condition of a fixed initial value of interference characteristic function which is

$$f(u_c) = J_k(u_c). \tag{20}$$

In the high-dimensional feature distribution space, after $k-1$ iteration, $k \ge 1$, the grey-scale sequence of music teaching evaluation should conform to the constraints of $J_k < \lambda_r$, and quantitative recursive analysis is used to obtain the DD values of the output index distribution data stream of music teaching evaluation:

$$R_k = \mathrm{Sim}(\alpha, d_i)\beta(d_i, J_k). \tag{21}$$

At this time, the information flow of music teaching evaluation is divided into K submatrices, which are (α, d_i) submatrices. Through clustering and merging the index parameters, the corresponding teaching resource allocation plan is made, and the efficient and accurate music teaching evaluation goal is achieved.

4. Experimental Verification

In order to verify the overall effectiveness of music teaching evaluation method based on data mining, the experiment platform MATLAB R2019B is developed, which is simulated in the environment of dominant frequency 1.

4.1. Experimental Preparation. In the datasets of Amazon Web Services (AWS), 1000 music data are randomly selected, of which 800 music data are used for training and 200 music data are used for testing to test the time of teaching evaluation, accuracy of evaluation results, and indicators of music type in the platform of music teaching system.

4.2. Experimental Results and Analysis. In this section, the experimental results and analysis have been discussed. Furthermore, the time analysis of teaching evaluation is discussed succeeding the accuracy of the results which are analyzed, and music type indicators are analyzed.

4.2.1. Time Analysis of Teaching Evaluation. Randomly selected 150 data of different song types for testing as experimental subjects. The superiority of this method is verified by comparing the evaluation time of this method with that of literature [3] method and literature [4] method. The result is shown in Figure 1.

It is shown from Figure 1 that by defining the mining object and target category based on the fuzzy association support degree of the set of music teaching data, the suggested technique can increase the effect of music teaching.

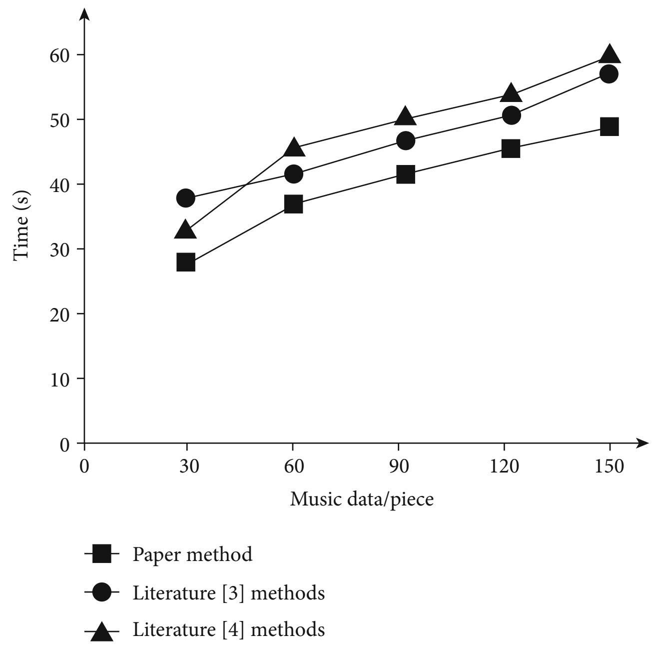


Figure 1: Results of teaching evaluation time.

Therefore, the time of music teaching evaluation is short, and only 47 s can effectively identify 150 different types of music data, and the efficiency is obviously higher than that of the other two methods.

4.2.2. Accuracy Analysis of Evaluation Results. By means of iteration, the higher the value of ω_n and ω_n in the analytical formula (15) is, the more accurate the evaluation result is and the higher the application value of evaluation method is. The calculation formula is

$$\omega_n = \frac{F}{m} \times 100\%. \tag{22}$$

Among them, m is the number of music teaching objectives, and F is already evaluated music teaching data. The accuracy of evaluation results is taken as the test index, and the test results of 3 different teaching evaluation methods are shown in Figure 2.

As can be seen from Figure 2, the ω_n value is above 91, and the ω_n value fluctuates from 83 to 90 when the literature [3] method and the literature [4] method are used to evaluate the music teaching. The evaluation findings acquired by this approach are more accurate than the testing results provided by the above methods, because this method uses a clustering algorithm to increase the clustering statistical ability.

From Figure 2, we can see that the values of music teaching by this method are all above 91. The values of the literature [3] method and literature [4] method fluctuate from 83 to 90. Compared with the testing results of the above methods, the evaluation results obtained by this method are more accurate, because this method makes use of clustering algorithm to improve the clustering statistical ability.

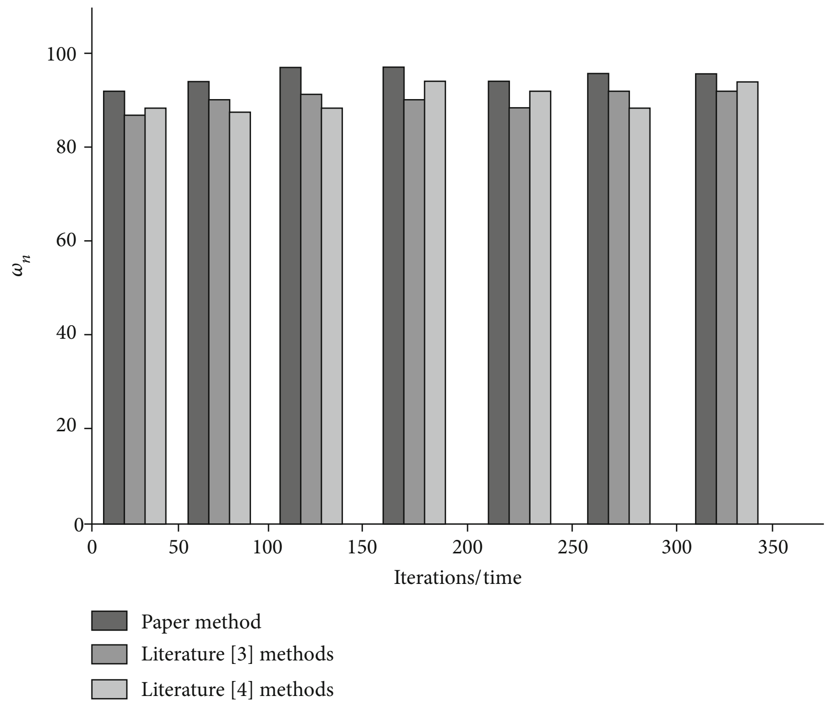
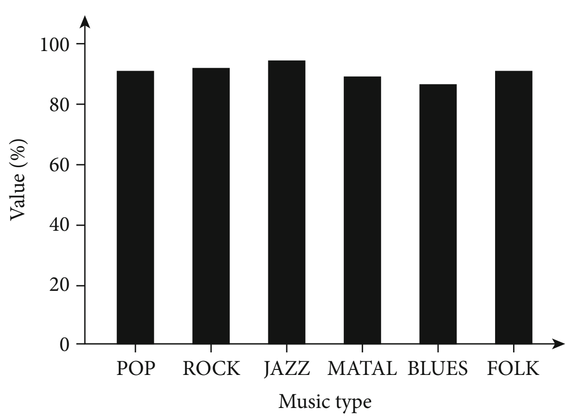
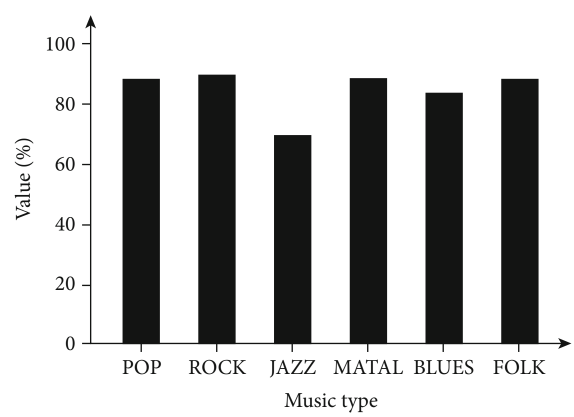


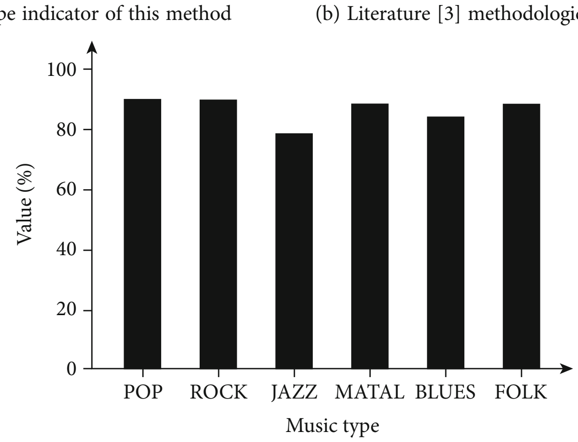
Figure 2: Assessment accuracy results.



(a) Results of the music type indicator of this method



(b) Literature [3] methodological musical type indicator results



(c) Literature [4] methodological musical type indicator results

Figure 3: Music genre metric results.

4.2.3. Analysis of Music Type Indicators. This study analyzes the outcomes of the analysis of music type indicators between this approach, literature [3] method, and literature [4] method, using 200 music data as the subject of music type analysis as shown in Figure 3.

Figure 3 shows that this method can effectively identify pop, rock, jazz, metal, blues, folk, and other music type indicators, but the literature [3] method and the literature [4] method cannot identify JAZZ type indicators effectively, because jazz music has a more dynamic rhythm, it requires linked fusion in a high-dimensional space in order to generate a clear impression of music type markers.

5. Conclusion

To summarize, musical data mining is utilized on the platform of the music teaching system to evaluate the time of teaching assessment, the accuracy of evaluation results, and indicators of music type. It can be found that data mining evaluation in the field of music teaching has significant results. The experiment was performed on a data that has been collected from a web source. The data has been randomly chosen and experimented by different algorithms performed in the previous studies such as teaching evaluation of time analysis, accuracy analysis, and indicators of music type. The outcomes obtained by the study has been compared with the different studies performed before. In the result comparison, the score achieved was between 83 and 90 for time evaluation analysis and 83 and 91 for data accuracy analysis. The results of the experiments suggest that the data mining method can be used to evaluate music training and that it has the advantages of a quick evaluation time, high accuracy, and unambiguous indicators. Although this paper has made some research results and has some significance for the evaluation of music teaching quality, however, there are still some deficiencies, which need to be further improved in the future.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that he has no conflict of interest.

References

[1] G. U. O. Xin, Z. H. A. N. G. Ming-huan, W. U. Liang, and C. H. E. N. Ying, “Research and implementation of comprehensive evaluation algorithms# br# for teaching quality in colleges and universities,” *Computer and Modernization*, vol. 6, p. 121, 2019.

[2] Y. Wang, Y. Ma, and S. Yu, “Reform of computer course scoring model for sergeant cadet education and its effect on teaching quality,” *Computer Engineering Science*, vol. 41, no. 1, pp. 229–233, 2019.

[3] X. Xia and J. Yan, “Construction of music teaching evaluation model based on weighted Naïve Bayes,” *Scientific Programming*, vol. 2021, Article ID 7196197, 9 pages, 2021.

[4] M. J. Espigares-Pinazo, J. M. Bautista-Vallejo, and M. García-Carmona, “Evaluations in the Moodle-mediated music teaching-learning environment,” *Technology, Knowledge and Learning*, vol. 27, no. 1, pp. 17–31, 2022.

[5] L. Yu, Y. Duan, and K. C. Li, “A real-world service mashup platform based on data integration, information synthesis, and knowledge fusion,” *Connection Science*, vol. 33, no. 3, pp. 463–481, 2021.

[6] D. Q. Tu, A. S. M. Kayes, W. Rahayu, and K. Nguyen, “IoT streaming data integration from multiple sources,” *Computing*, vol. 102, no. 10, pp. 2299–2329, 2020.

[7] J. Hu, J. Fang, Y. Du, Z. Liu, and P. Ji, “Application of PLS algorithm in discriminant analysis in multidimensional data mining,” *The Journal of Supercomputing*, vol. 75, no. 9, pp. 6004–6020, 2019.

[8] N. Kousika and K. Premalatha, “An improved privacy-preserving data mining technique using singular value decomposition with three-dimensional rotation data perturbation,” *The Journal of Supercomputing*, vol. 77, no. 9, pp. 10003–10011, 2021.

[9] S. Singh and S. Srivastava, “Kernel fuzzy C-means clustering with teaching learning based optimization algorithm (TLBO-KFCM),” *Journal of Intelligent & Fuzzy Systems*, vol. 42, no. 2, pp. 1051–1059, 2022.

[10] C. Zhang, M. Li, H. Wang, and N. Wang, “Auxiliary decision support model of sports training based on association rules,” *Mobile Information Systems*, vol. 2021, 8 pages, 2021.

[11] C. Wang, Y. L. Gao, X. Z. Kong, J. X. Liu, and C. H. Zheng, “Unsupervised cluster analysis and gene marker extraction of scRNA-seq data based on non-negative matrix factorization,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, pp. 458–467, 2022.

[12] A. Onovo, A. Atobatele, A. Kalaiwo et al., “Aggregating loss to follow-up behaviour in people living with HIV on ART: a cluster analysis using unsupervised machine learning algorithm in R,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 458–467, 2022.

[13] A. Fatikhurrizqi and A. W. Wijayanto, “Comparison of regional cluster analysis according to inclusive development indicators in java island 2018 between hierarchical and partitioning clustering strategies,” *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 6, no. 2, pp. 143–150, 2021.

[14] S. Ghosh, D. V. Thang, S. C. Satapathy, and S. N. Mohanty, “Fuzzy rule based cluster analysis to segment consumers’ preferences to eco and non-eco friendly products,” *International Journal of Knowledge-based and Intelligent Engineering Systems*, vol. 24, no. 4, pp. 343–351, 2021.

[15] Z. Bin, “Regional enterprise economic development dimensions based on k-means cluster analysis and nearest neighbor discriminant,” *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 6, pp. 7365–7375, 2020.