

Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Data Mining Techniques for the Preservation and Inheritance of Classical Vocal Music in Modern Society

Xiaoqing Chi^{1,†}

1. Conservatory of Music, Henan University of Economics and Law, Zhengzhou, Henan, 450000, China.

Submission Info

Communicated by Z. Sabir
Received November 17, 2024
Accepted February 20, 2025
Available online March 17, 2025

Abstract

This study firstly describes the application of data mining technology in the protection and inheritance of classical vocal music, and secondly proposes the basic steps of association rule algorithms in data mining technology for exploring the association between classical vocal music and user collections. This study selects part of the data of a music platform, scientifically analyzes the relationship between user collection and classical vocal works, and empirically proves the classical vocal music protection and inheritance method based on data mining technology proposed in this paper. The study concludes that the probability of users collecting the Ninth Symphony and Carmen is 63.122%, and the probability of collecting the Ninth Symphony and then Carmen is 86.544% when the Ninth Symphony is collected. The experimental data is a strong evidence for the preservation and inheritance of classical vocal music in modern society.

Keywords: Data mining techniques; Association rules; Classical vocal music; Preservation and heritage.

AMS 2010 codes: 03B70

[†]Corresponding author.

Email address: zhangchen0605@163.com

1 Introduction

On the road to the development of Chinese folk music, the results achieved by borrowing from the West are promising. Although the entry of western culture into China has caused a certain impact on Chinese culture, it doesn't mean that only China's own local culture needs to be developed; culture is more valuable and meaningful only after mutual exchanges and borrowing. In order to obtain better protection and inheritance of classical vocal music in modern society, it is necessary to realize the self-development after borrowing, which is the real meaning of borrowing [1-4]. Music is not like medicine, mathematics, according to the book is no vitality, China's fifty-six nationalities of ethnic music and cultural resources are extremely rich, these music and culture is the cornerstone of the survival of classical vocal music. Chinese culture shows a kind of simplicity and innocence, no matter in the academic or artistic field, the Chinese people strive to use the simplest and the least amount of words to express more content, and create the most moving flavor with the least amount of ink and pen [5-6]. Chinese classical vocal music, which is built on the foundation of traditional culture, uses folk music as the material and thus develops continuously. The creation of folk vocal music works is based on traditional culture, using western composition techniques, and constantly exploring and perfecting. And nowadays, the creation of modern folk vocal music has done better in these aspects. Chinese art is mostly linear art, emphasizing the horizontal beauty of melodic lines, while Western art is mostly about vertical three-dimensional beauty [7-9]. To borrow western music as the only feasible idea and theory in the creation and performance of folk vocal music is itself against scientific thought and biased practice. How to rationally integrate the Internet resources to reduce the additional cost of protection and inheritance, how to maximize the mining of data and effectively apply it to the protection and inheritance of classical vocal music is an urgent problem for the majority of researchers [10-12]. The use of cloud computing technology can further optimize the allocation of resources, the direct use of data resources stored in the cloud, which can minimize the pressure on the end customer, and promote the continuous growth of data stored in the cloud. Data mining generally refers to a process of searching for information hidden in massive data through algorithms [13-14]. As data mining technology is more and more widely used in various industries and trades, research workers classify different research objects, construct the research model corresponding to each class, and collect a series of user data, and then utilize a variety of different types of algorithms or relational graphical models or classifiers of music resources, and finally build a music expert system based on big data. It helps vocal music protection and inheritance, and contributes to the continuation and development of vocal music culture [15-16].

The author collects the relevant data of a music platform, conducts a big data analysis of the playback and dissemination of classical vocal music on a music platform, and then analyzes the relationship between users' collection behavior and classical vocal music works by using the Apriori algorithm, and further analyzes the data on the users' feelings, so as to understand the users' views and feelings about classical vocal music works, thus further promoting the protection and inheritance of classical vocal music in modern society. The study will further analyze the relationship between users' collection behavior and classic vocal works, and further analyze the data of users' feelings, so as to understand users' views and feelings towards classic vocal works, and thus further promote the protection and inheritance of classic vocal music in modern society.

2 Method

2.1 Application of Data Mining Technology in the Protection and Inheritance of Classical Vocal Music

In recent years in the classic vocal music platforms and network platforms have most people click on the songs and videos of the amount of data gradually increased, which makes the music network platforms not only get rapid development, but also the listener's favorite songs for feedback analysis. For major music platforms, the classic sound resources and user experience are important indicators to reflect the good or bad of a platform. The classic vocal resources and user experience can be analyzed and organized by data mining, on the one hand, it can better analyze the most popular songs and improve the overall click rate of the classic vocal platform, on the other hand, through the listener often listen to the type of songs and singers to analyze the data, and better provide the listener with songs and singers of similar and similar types of music styles. Through the application of data mining technology in the field of classical vocal music, it not only promotes the overall development of online music platforms, but also brings a very lucrative source of funding for individuals and economic companies, thus promoting the development of classical vocal music in a better and more effective way.

2.2 Data Mining Theory

Knowledge discovery is a newly emerged term with information technology, knowledge explosion era, it is from a large amount of incomplete, fuzzy, noisy data to distill regular, practical and strong knowledge and information, which people can use to improve the workflow, enhance the enterprise system and increase the efficiency of the enterprise. Data mining is the core part of knowledge discovery, which is more specialized than knowledge discovery, and can be said to be an advanced stage of using knowledge to accumulate data, and its main function is to use various algorithms to find out its intrinsic laws and patterns from a large amount of data, and to assist managers to make effective decisions. Data mining is also known as knowledge discovery in the database, it is with the help of advanced technology (mining software), a large amount of knowledge and methods (customer segmentation methods), some shallow, rough, messy information for cleaning and conversion, through the organization of the data after the discovery of the potential laws and connections, the development of future things to make an effective prediction, guiding the management staff to examine the current state of enterprise development, and make timely evaluation and corresponding decision-making in a timely manner. Data mining is an advanced technology to deal with a large amount of data, it is wider in scope and easier to use, it is in the absence of clear assumptions, you can discover knowledge, mining information, get with previously unknown but effective, usable information, so as to achieve the goal of saving resources, improve revenue to increase income, so that the enterprise is in a more advantageous competitive position.

2.3 Association rules

2.3.1 Definition of association rules

Association analysis, also known as "shopping basket" analysis, is mainly used to determine the connection between different domains in the data, to find out the dependency relationship between multiple domains, is a more important method in data mining. Association rule is to find two or more data items between the values of a certain regularity, such as a certain law is called association. We can develop new strategies based on these laws discovered. Totaling out the original unknown

association rules can promote business development. The purpose of association analysis is to uncover the hidden network of association rules in the database. There are many types of association rules in large databases, such as simple association, temporal association, and causal association. Most association rules always have to consider certain parameters due to their practicality, scientific validity, and successful use. [17] The parameters to be considered for association rules in general are the number of valid items, support, credibility, etc. By limiting these parameters, the diggers can discover rules that are more compatible with the specific requirements.

Let the association rule mining in a transaction number mining library can be described as follows:

Let $I = \{i_1, i_2, \dots, i_m\}$ be a collection of items, and the transaction database $D = \{t_1, t_2, \dots, t_n\}$ be composed of a series of transactions with a unique identifier TID, and each transaction $t_i (i = 1, 2, \dots, n)$ corresponds to a subset on I . Then the association rule mining in a thing database has the following special rules:

- 1) Let $I_1 \subseteq I$, the support of itemset I_1 on dataset D is the percentage of transactions containing I_1 in D , i.e., $Support(I_1) = ||t \in D | I_1 \subseteq t || / ||D||$.
- 2) For itemset I and transaction database D , all itemsets that satisfy the user-specified minimum support, i.e., a non-empty subset greater than or equal to I of Minsupport, are called frequent itemsets or large itemsets. A frequent itemset that picks out all frequent itemsets that are not contained by other itemsets is called a maximum frequent itemset or a maximal itemset.
- 3) An association rule of the form $I_1 \Rightarrow I_2$ defined on I and D is given by satisfying a certain level of confidence or then confidence, the so called confidence level of the rule is the ratio of the number of transactions containing I_1 and I_2 to the number of transactions containing I_1 , i.e.:

$$Confidence(I_1 \Rightarrow I_2) = \text{sup port}(I_1 \cup I_2) / \text{sup port}(I_1) \quad (1)$$

where $I_1, I_2 \subseteq I, I_1 \cap I_2 = \emptyset$.

- 4) An association rule that D satisfies minimum support and minimum confidence on I is called a strong association rule. The association rules usually described are the strong association rules.

In general, given a transactional database, the association rule mining problem is the process of finding strong association rules by specifying minimum support and confidence picks. The problem of mining association rules can be broken down into two subproblems. The first is to find frequent itemsets, through the minimum support given by the user, to find all frequent itemsets, i.e., to meet the Support is not less than the Minsupport of all subsets of items. The second task is to generate association rules that find association rules with confidence values not lower than Minconfidence in each most frequent item set according to the minimum confidence provided by the user.

2.3.2 Apriori algorithm

The Apriori algorithm incrementally discovers frequent itemsets by growing the number of item elements. First 1-frequent itemsets L_1 are generated, followed by 2-frequent itemsets L_2 until the algorithm stops when the number of frequent itemset elements cannot be expanded any further. In the k rd loop, the process first produces a collection of dust k candidate itemsets C_k and then generates k – frequent itemsets L_k by scanning the database to generate support and testing [18].

A sample transaction database is given below and the Apriori algorithm is implemented on it. The sample transaction database is shown in Table 1.

Table 1. Sample transaction database

TID	Itemset	TID	Itemset
1	A,B,C,D	4	B,D,E
2	B,C,E	5	A,B,C,D
3	A,B,C,E		

Trace the execution of the Apriori algorithm for the transaction database shown in Table 1

1) L_1 Generation

Their support numbers by scanning the database:

$$C_1 = \{(A, 3), (B, 5), (C, 4), (D, 3), (E, 3)\} \quad (2)$$

Pick $\text{minsup_count} \geq 2$ the set of items $L_1 = \{A, B, C, D, E\}$.

2) L_2 generate

The 2-candidate sets are generated by L_1 and their support numbers are obtained by scanning the database with the number of elements C_s^2 :

$$C_2 = \{(AB, 3), (AC, 3), (AD, 2), (AE, 1), (BC, 4), (BD, 3), (BE, 3), (CD, 2), (CE, 2), (DE, 1)\} \quad (3)$$

The set of items selected for $\text{minsup_count} \geq 2$ forms the set of 2-frequent items:

$$L_2 = \{AB, AC, AD, BC, BD, BE, CD, CE\} \quad (4)$$

3) L_3 -generate

By L_2 to generate 3-candidate sets and get their support numbers by scanning the database:

$$C_3 = \{(ABC, 3), (ABD, 2), (ABE, 1), (ACD, 2), (ACE, 1), (BCD, 2), (BCE, 2), (BDE, 1), (CDE, 0)\} \quad (5)$$

The set of items selected for $\text{minsup_count} \geq 2$ forms a 3-frequent itemset:

$$L_3 = \{ABC, ABD, ACD, BCD, BCE\} \quad (6)$$

4) L_4 -generate

By L_3 to generate 3-candidate sets and get their support numbers by scanning the database:

$$C_4 = \{(ABCD, 2), (ABCE, 1), (BCDE, 0)\} \quad (7)$$

The set of items selected $\text{minsup_count} \geq 2$ form a 4-frequent itemset $L_4 = \{ABCD\}$.

5) L_5 generates

By L_4 to generate 5-candidate set $C_5 = \emptyset$ and the algorithm stops. Thus all frequent itemsets are $\{A, B, C, D, E, AB, AC, AD, BC, BD, BE, CD, CE, ABC, ABD, ACD, BCE, ABCD\}$.

Also, it is easy to get the maximum frequent itemset as $\{ABCD, BCE\}$.

6) Association rule generation

Find the association rules whose Confidence is not less than Minconfidence in each maximal frequent itemset by the minimum confidence level plotted by the user:

$$\text{confidence}(A \Rightarrow BCD) = \frac{2}{3} \quad \text{confidence}(B \Rightarrow ACD) = \frac{2}{5} \quad \text{confidence}(C \Rightarrow ABD) = \frac{2}{4}$$

$$\text{confidence}(D \Rightarrow ABC) = \frac{2}{3} \quad \text{confidence}(AB \Rightarrow CD) = \frac{2}{3} \quad \text{confidence}(AC \Rightarrow BD) = \frac{2}{3}$$

$$\text{confidence}(AD \Rightarrow BC) = 1 \quad \text{confidence}(BC \Rightarrow AD) = \frac{2}{4} \quad \text{confidence}(BD \Rightarrow AC) = \frac{2}{3}$$

$$\text{confidence}(CD \Rightarrow AB) = 1 \quad \text{confidence}(ABC \Rightarrow D) = \frac{2}{3} \quad \text{confidence}(ABD \Rightarrow C) = 1$$

$$\text{confidence}(ACD \Rightarrow B) = 1 \quad \text{confidence}(BCD \Rightarrow A) = 1 \quad \text{confidence}(B \Rightarrow CE) = \frac{2}{5}$$

$$\text{confidence}(C \Rightarrow BE) = \frac{2}{4} \quad \text{confidence}(E \Rightarrow BC) = \frac{2}{3} \quad \text{confidence}(BC \Rightarrow E) = \frac{2}{4}$$

$$\text{confidence}(BE \Rightarrow C) = \frac{2}{3} \quad \text{confidence}(CE \Rightarrow B) = 1.$$

out of minconfidence -70%, so born:
into a valid association rule for $AD \Rightarrow BC, CD \Rightarrow AB, ABD \Rightarrow C, BCD \Rightarrow A, CE \Rightarrow B$.

Apriori, as the classic frequent itemset dusting algorithm, has a landmark role in database mining. However, with further research, its shortcomings are also exposed. The Apriori algorithm has two fatal performance bottlenecks. One is scanning the transaction database many times, which requires a large input/output load, e.g., if a frequent large itemset contains 10 items, then it is necessary to scan the transaction database at least 10 times: the second is the possibility of generating a huge amount of candidate sets, e.g., HL_{i-1} generates k-

candidate sets C_i is exponentially growing, 10^4 1-frequent itemsets may generate a 2-candidate set of close to 107 elements, and such large candidate sets are a big problem to the both time and main memory space. Because of this, many scholars, including Agrawal, have proposed improvements to the Apriori algorithm, such as data partitioning based approach, hashing based approach, sampling based approach, Close algorithm, FP-tree algorithm, etc. Since the efficiency and storage issues of the algorithm are not examined in this paper, Apriori algorithm is applied in this paper to mine the consumption behavior of mobile customers [19].

3 Results and discussion

3.1 Big Data Analysis of Internet Dissemination of Classical Vocal Music

The experiments in this section first crawl the data of twelve classical vocal pieces on a classical vocal music playback platform using a web crawler. Next, the top ten tracks of each genre and their playback volume are selected by sorting them according to playback volume. Then, the Internet correlations between song genres are quantitatively calculated. Finally, the correlations between song genres on the Internet are adjusted to account for the Internet dissemination heat. If both genre A and genre B have audio or video transmission of track A in the Internet, it is considered that genre A and genre B have some Internet relevance for track A, and their relevance is set to 1, otherwise it is 0. The relevance of the top ten tracks of each genre is accumulated to determine the overall relevance. The Internet popularity of each genre is measured by the cumulative relative number of views for the top ten tracks. The overall Internet relevance matrix for classical vocal works is shown in Table 2.

Table 2. The overall correlation degree matrix of the vocal music works network

	Symphony No. 9	Magic Flute	Aida	Turandot	Carmen	Cruciform	Messiah	The Barber Of Sevilla	Peras And Melinda	Spring Sacrifice	The Ring Of Niberlong	Bogie And Beth
Symphony No. 9	12	8	7	10	13	4	8	3	5	1	8	0
Magic Flute	8	12	2	2	8	3	9	2	0	0	2	3
Aida	1	7	8	12	3	5	13	4	4	3	2	4
Turandot	7	1	5	10	3	4	8	0	4	0	0	2
Carmen	5	5	0	2	8	7	4	3	0	0	0	1
Cruciform	6	7	3	6	6	8	5	1	0	1	1	0
Messiah	3	8	1	4	7	2	11	10	2	1	0	1
The Barber Of Sevilla	1	6	4	1	1	6	4	9	0	2	6	1
Peras And Melinda	13	2	6	9	7	0	1	0	9	3	6	1
Spring Sacrifice	4	11	3	1	0	1	6	7	0	7	2	1
The Ring Of Niberlong	4	0	1	1	1	2	0	2	1	0	11	0
Bogie And Beth	2	6	4	3	0	9	8	3	0	0	1	10

Since the top ten tracks of classic vocal music have a huge advantage in terms of Internet dissemination heat, they are ignored for the time being for the sake of data fitting. The relationship between the overall relevance of classic vocal works and the playback quality is shown in Figure 1.

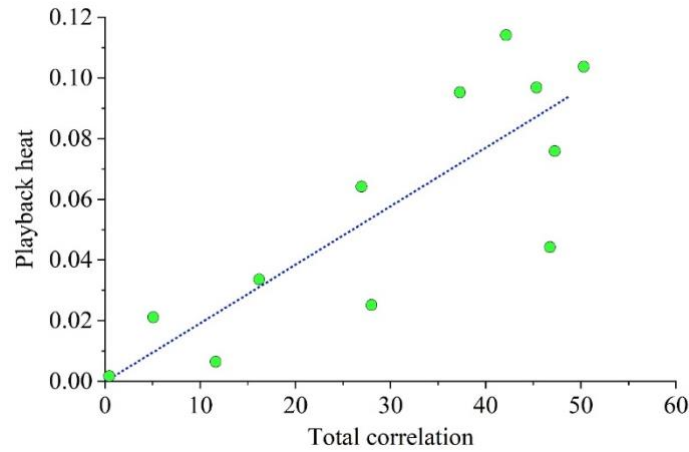


Figure 1. The relationship between the general correlation and the heat of the playback

Taking into account the effect of track playback on relevance (more played tracks correspond to higher relevance), the calculation of weighted relevance is introduced. The weighted relevance matrix for the network of classical vocal works is shown in Table 3.

Table 3. The classical vocal music works network weighted correlation matrix

	Symphony No. 9	Magic Flute	Aida	Turandot	Carmen	Cruciform	Messiah	The Barber Of Sevilla	Peras And Melinda	Spring Sacrifice	The Ring Of Niberlong	Bogie And Beth
Symphony No. 9	0.047	0.043	0.031	0.051	0.048	0.03	0.023	0.002	0.009	0.008	0.032	0.009
Magic Flute	0.101	0.345	0.115	0.142	0.161	0.201	0.153	0.134	0.006	0.063	0.079	0.004
Aida	0.072	0.087	0.074	0.097	0.053	0.042	0.036	0.022	0.03	0.017	0.03	0.018
Turandot	0.027	0.04	0.06	0.082	0.014	0.043	0.044	0.028	0.004	0	0.021	0.012
Carmen	0.073	0.075	0.029	0.027	0.074	0.012	0.034	0.016	0.006	0.029	0.023	0.008
Cruciform	0.041	0.061	0.043	0.064	0.051	0.105	0.035	0.025	0.015	0.011	0.028	0.011
Messiah	0.016	0.087	0.02	0.051	0.033	0.038	0.106	0.068	-0.001	-0.004	0.014	-0.004
The Barber Of Sevilla	0.01	0.035	0.02	0.025	0.003	0.046	0.031	0.075	0.002	0.008	0.008	0.007
Peras And Melinda	0.004	0.001	0.007	0.003	0.003	0.01	0.001	0.003	0.009	0.003	0.008	0
Spring Sacrifice	0.019	0.033	0.028	0.009	0.02	0.009	0.023	0.007	0.014	0.038	0.016	0.02
The Ring Of Niberlong	0.008	0.008	0.003	0.004	0.002	0.008	0.003	0.009	0.016	0.004	0.001	0.01
Bogie And Beth	0.002	0.007	0.006	0.002	0.006	0.014	0.001	0.002	0.013	0.001	0.005	0.016

The relationship between the web-weighted relevance of classical vocal works and playback heat is shown in Figure 2.

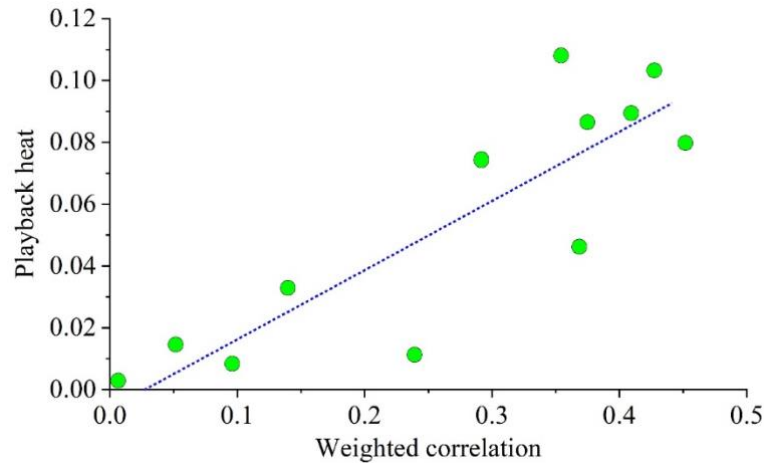


Figure 2. The relationship between weighted correlation and playback

A music platform is used to do data mining of the world's most influential and representative classical vocal works. Firstly, the 12 selected classic vocal works and the music platform song list are searched (based on the data of June 14, 2023), and then the irrelevant information is filtered out, and finally the statistics are made. The search frequency of classic vocal music and related statistics are shown in Table 4. It can be seen that there is a huge gap between the 12 classic vocal works in terms of comment heat. At the same time, the data also shows that "Carmen" constitutes a significant advantage over other classic vocal works in terms of comment heat. Its comment volume is more than 4000, and none of the other classic vocal works have tracks with more than 4000 comments.

Table 4. Classic vocal search frequency and related statistics

Name	Word Frequency	Top Number Of Comments	Average Number Of Comments
Symphony No. 9	9945	3512	245
Magic Flute	5690	1015	185
Aida	4671	2455	65
Turandot	6680	86	48
Carmen	3433	48532	1732
Cruciform	3632	1226	182
Messiah	6607	988	89
The Barber Of Sevilla	7611	745	1533
Peras And Melinda	8297	3221	1205
Spring Sacrifice	8582	2896	125
The Ring Of Niberlong	8255	1633	212
Bogie And Beth	4139	1874	91

3.2 Application of Apriori algorithm based association rules in the protection and inheritance of classical vocal music

Taking the data of a classical vocal music playback platform as an example, when users collect songs, they are faced with a huge amount of information on many classical vocal works, and often show dazzling, unable to quickly find a satisfactory classical vocal work, which increases the time and reduces the experience of vocal music appreciation. Users' preferences and tastes for classical vocal

works are generally fixed, and there is a certain rule of choosing different classical vocal works to match according to different types of classical vocal works. Classic vocal works are interconnected, while some classic vocal works are in opposition or competition (negative correlation), these laws are hidden in a large amount of historical data, if you can discover the laws of users' preferences for classic vocal works through data mining, you can quickly identify the taste of music users. The data related to classic vocal works collected by collectors are obtained from it. The collection data is shown in Table 5.

Table 5. Collect data

Collectors	Symphony No. 9	Magic Flute	Aida	Turandot	Carmen	Cruciform	Messiah	The Barber Of Sevilla	Peras And Melinda	Spring Sacrifice	The Ring Of Niberlong	Bogie And Beth
Collectors 1	0	1	1	1	1	1	1	1	1	1	0	1
Collectors 2	1	1	0	1	1	0	1	1	0	1	0	0
Collectors 3	1	0	1	1	1	0	1	0	0	0	1	0
Collectors 4	1	1	0	1	1	1	1	1	0	0	0	0
Collectors 5	1	1	1	0	1	1	0	0	0	1	0	0
Collectors 6	1	1	0	1	1	1	1	0	1	0	1	1
Collectors 7	1	1	1	0	1	1	1	0	1	0	1	0
Collectors 8	0	0	1	1	1	1	0	1	0	1	1	0
Collectors 9	0	0	1	0	0	0	1	0	1	1	0	0
Collectors 10	1	0	1	1	1	1	0	0	0	1	0	0
Collectors 11	1	0	0	0	0	1	1	1	0	1	0	0
Collectors 12	0	1	0	1	0	1	1	1	1	1	0	1
Collectors 13	1	0	1	1	1	0	1	1	1	0	0	1
Collectors 14	1	0	1	0	0	0	0	1	1	0	1	0
Collectors 15	1	1	0	1	1	0	0	0	1	0	0	0
Collectors 16	1	1	1	0	0	0	1	1	0	1	0	0
Collectors 17	1	0	0	0	1	0	1	0	0	1	1	1
Collectors 18	1	1	1	1	1	0	1	1	1	0	1	0
Collectors 19	1	0	1	0	0	0	0	1	1	0	1	1
Collectors 20	1	1	1	0	1	1	0	0	0	0	0	0

The minimum support of the above rules was set to 0.2 and the minimum confidence level was set to 0.65. 7 association rule results were obtained and the association rule results are shown in Table 6.

Table 6. Association rule results

Foreterm	Afterterm	Support%	Confidence%
Symphony No. 9	Carmen	63.122	86.544
Magic Flute	Aida	53.592	76
The Barber Of Sevilla	The Ring Of Niberlong	50	73.651
Aida	Bogie And Beth	53.641	65.951
Cruciform	Magic Flute	54.955	66.653
Turandot	Symphony No. 9	54.875	62.152
Carmen	Cruciform	60.253	72.541

Rule 1: The probability of having collected Symphony No. 9 and Carmen is 63.122%, and the probability of having collected Symphony No. 9 and then Carmen is 86.544%.

Rule 2: The probability of having collected The Magic Flute and Aida is 53.592%, and the probability of having collected The Magic Flute and then Aida is 76%.

Rule 3: The probability of having collected The Barber of Seville and The Ring of the Nibelungen is 50.0%, and the probability of having collected The Barber of Seville and then The Ring of the Nibelungen is 73.651%.

Rule 4: The probability of having collected Aida and Porgy and Bess is 53.641%, and the probability of having collected Aida and then Porgy and Bess is 65.951%.

Rule 5: The probability of having collected Matthew's Passion and The Magic Flute is 54.955%, and the probability of having collected Matthew's Passion and then The Magic Flute is 66.653%.

Rule 6: The probability of having collected Turandot and Symphony No. 9 is 54.875%, and the probability of having collected Turandot and then Symphony No. 9 is 62.152%.

Rule 7: The probability of having collected Carmen and Matthew's Passion is 60.253%, and the probability of having collected Carmen and then Matthew's Passion is 72.541%.

Therefore, synthesizing the results of the above research, a classical vocal music playback platform can make use of big data to recommend music to users who like to listen to classical vocal works and collect classical vocal works, so that users can hear more classical vocal works.

3.3 Analysis of the effect of classical vocal music protection and inheritance

The collected sentiments of a classical vocal music playing platform were processed by word division using Phthon software to filter out the top 50 words, of which learning, culture, China, tradition and inheritance were the top five words with the highest frequency of occurrence. Through the analysis, it was concluded that most users showed strong interest in classical vocal music and believed that by listening to classical vocal music, they not only learned more knowledge about national culture, but also improved their music literacy and artistic expression. Some users have even started to compose some works. The top 50 words are shown in Table 7.

Table 7. Top 50 vocabulary

Serial Number	Word	Frequency	Serial Number	Word	Frequency
1	Learning	490	26	Innovate	70
2	Culture	475	27	Deep	68
3	Vocal Music	428	28	Charm	64
4	Tradition	399	29	Peoples	63
5	Pass On	385	30	Language	61
6	China	369	31	Region	58
7	Peoples	345	32	Connotation	58
8	Art	336	33	Place	57
9	Develop	236	34	Tune	54
10	Native	223	35	Rhythm	54
11	Affections	200	36	Disseminate	54
12	History	182	37	Value	54
13	Platform	136	38	Country	52
14	Form	132	39	Cultural Heritage	51
15	Appreciation	128	40	Spirit	51
16	Uniqueness	115	41	Excellence	50
17	Region	108	42	Experience	48
18	Folk	92	43	Elegance	48
19	Feeling	86	44	National Music	47
20	Understand	80	45	Interest	46
21	Feature	80	46	Primitive Ecology	44
22	Practice	78	47	Chinese Nation	43
23	Lyrics	75	48	Love	41
24	To Create	74	49	Singing	40
25	Protect	74	50	Protection And Inheritance	35

4 Conclusion

This study analyzes the protection and inheritance of modern vocal works using association rule algorithms, and obtains some conclusions that play a role in promoting the protection and inheritance of classic vocal works. In the big data analysis of the performance of classic vocal works in a music platform, it is found that “Carmen” has a significant advantage in the comment heat its comment volume exceeds 4000. No other classical vocal works have more than 4000 comments. In the analysis of the effect of the protection and inheritance of classical vocal music, the top 50 words in the ranking of the user’s feelings were screened out, of which the top five words with the highest frequency were learning, culture, China, tradition and inheritance. Based on this, it can be concluded that the majority of users on this music platform have a strong interest in classical vocal music.

References

- [1] Kuang, J., & He, L. (2022). From oblivion to reappearance: A multi-faceted evaluation of the sustainability of folk music in Yunnan province of China. *Sage Open*, 12(3), 21582440221117806.

- [2] Jincheng, X., Pattananon, N., & Mei, L. S. (2023). “Yi” Music and the Teaching Methods of Folk Music in China. *Journal of Modern Learning Development*, 8(6), 213-221.
- [3] Ping, S., Chuangprakhon, S., Santaveesuk, P., & You, L. (2024). The Evolution of Dong Small Songs and Cultural Change in Chinese Folk Music. *Journal of Ecohumanism*, 3(3), 1530-1540.
- [4] Li, R. (2022). Chinese folk music: Study and dissemination through online learning courses. *Education and Information Technologies*, 27(7), 8997-9013.
- [5] Xu, N. (2018). Analysis of the Correlation Between Folk Music Education and Chinese Traditional Culture. *Educational Sciences: Theory & Practice*, 18(5).
- [6] Chen, G., Sensai, P., & Seeyo, W. (2023). Knowledge of the Historical Development of Chinese Xiansuo Folk Music in Qinghai Province’s Hehuang District. *Education Quarterly Reviews*, 6(1).
- [7] Wang, D. (2021). Review of the Book “Morphology of Chinese Folk Music”. *Open Access Library Journal*, 8(8), 1-7.
- [8] Wang, Q., & Luo, J. (2022). Chinese folk music and its effect on human mathematical thinking skills. *Interdisciplinary Science Reviews*, 47(1), 97-106.
- [9] Li, F. (2020, April). Aesthetic trend of folk vocal music under the influence of big data thinking. In 2020 International Conference on Big Data and Informatization Education (ICBDIE) (pp. 96-98). IEEE.
- [10] Yu, G. (2021, February). Folklore thinking on the inheritance and protection of Miao nationality folk songs based on computer. In *Journal of Physics: Conference Series* (Vol. 1744, No. 3, p. 032076). IOP Publishing.
- [11] Ding, L. (2023). Discussion on the Protection and Inheritance of Primitive Folk Songs. *Frontiers in Art Research*, 5(16).
- [12] Mengting, Y. (2018). ON THE STATIC PROTECTION AND DYNAMIC INHERITANCE OF CHINESE NATIONAL MUSIC. *The Scientific Heritage*, (30-1 (30)), 3-6.
- [13] Wang, S., & Lou, G. (2023, August). Data Construction of Music Culture Communication Based on Data Mining Technology. In *EAI International Conference, BigIoT-EDU* (pp. 193-200). Cham: Springer Nature Switzerland.
- [14] Ma, L. (2019). Traditional music protection system from the ecological perspective based on big data analysis. *Ekoloji Dergisi*, (107).
- [15] Wu, L., & Li, C. (2022, May). Analysis of the Dissemination Topology Network and Inheritance of Chinese Folk Music Culture based on the Deep Retrieval Algorithm of Internet Information. In 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 681-685). IEEE.
- [16] Yu, T., Wang, X., Xiao, X., & Yu, R. (2024, June). Harmonizing Tradition with Technology: Using AI in Traditional Music Preservation. In 2024 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
- [17] Haiyan Yao. (2024). The Two-Way Drive in Modern Informational Music Teaching and Traditional Music Culture Inheritance. *Applied Mathematics and Nonlinear Sciences*(1).
- [18] Yang Yang, Yu Ting Li, Yong Hua Huo, Zhi Peng Gao & Lan Lan Rui. (2024). Alarm Log Data Augmentation Algorithm Based on a GAN Model and Apriori. *Journal of Computer Science and Technology*(4), 951-966.
- [19] Beibei Wang. (2022). Automatic Integration Algorithm of Vocal Performance Learning Materials Based on Multidimensional Association Rules. *Mathematical Problems in Engineering*.