Hindawi
Mobile Information Systems
Volume 2022, Article ID 3983201, 9 pages
https://doi.org/10.1155/2022/3983201

*Research Article*

# Music Emotion Representation Learning Based on Multisource Data Fusion and Its Application

**Haibo Zhang**

*Nanchang Institute of Technology, Nanchang, Jiangxi 330044, China*

Correspondence should be addressed to Haibo Zhang; 28849631@qq.com

Received 25 May 2022; Revised 16 June 2022; Accepted 29 July 2022; Published 27 September 2022

Academic Editor: Praveen Kumar Reddy Maddikunta

Copyright © 2022 Haibo Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the multisource data fusion and fusion of music emotion expression learning and its application, this paper further analyzes the actual influence of music emotion and Internet multisource data in emotion learning. This paper uses multisource data to improve the professionalism and concentration of music emotion and uses modern Internet technology to help users quickly integrate into music emotion learning. At the same time, the multisource data structure can develop the music learning structure to a high-quality level. It is precisely because of the emphasis of the multisource data model architecture that the learning mechanism of musical emotion representation can be continuously updated and improved, which is mutually promoted jointly with Internet technology.

## 1. Introduction

Musical emotion is the basis of musical emotion. In order to better identify music emotions, help the service platform to perform precise positioning, and systematically divide music files in a timely manner, so as to achieve faster and more efficient personalized recommendation, it is necessary to use multisource data for structural sorting to meet user needs. At the same time, Internet technology can timely and systematically divide music files to achieve faster and more efficient personalized recommendations [1]. Users need to use multisource data to sort the structure to meet their needs. Multisource data are mainly composed of matrix sequences and multimodal structures. Through the integration and innovation of musical attributes and musical emotional expression, maximizing the musical emotional information can not only promote the accurate grasp of musical emotion but also comprehensively deal with the emotional level, screen out the music collection in line with the musical emotion, and improve the learning efficiency of musical emotion [2]. The research of musical emotion by Internet technology can accelerate the integration of musical emotion and push the musical emotion up to a higher expression level through the brewing of musical emotion. The Internet model based on user data information and music information can help users find their own music quickly and accurately [3]. The multivariate data structure is shown in Figure 1.

In the past, music emotion recognition was mainly based on the key data of audio for structure conversion. At the same time, there are music spectrum and music rhythm characteristics to classify music emotion. Therefore, in all kinds of music emotion analysis, music spectrum features are widely used in music emotion learning.The use of music spectrum features can analyze and manage the key audio information, making music emotion easier to express [4].

## 2. Multisource Data Model

*2.1. Data Model Types.* The multisource data model includes resolution data, multidimensional data, and various types of data [5]. These data are different in data form such as format, unit, resolution, accuracy, and internal characteristics such as attributes and content. Figure 2 shows the buffer of the audio data volume.

Digital map is the concentrated expression of multivariate data and the main body of multisource data. It mainly includes digital line drawing, digital grid map, digital

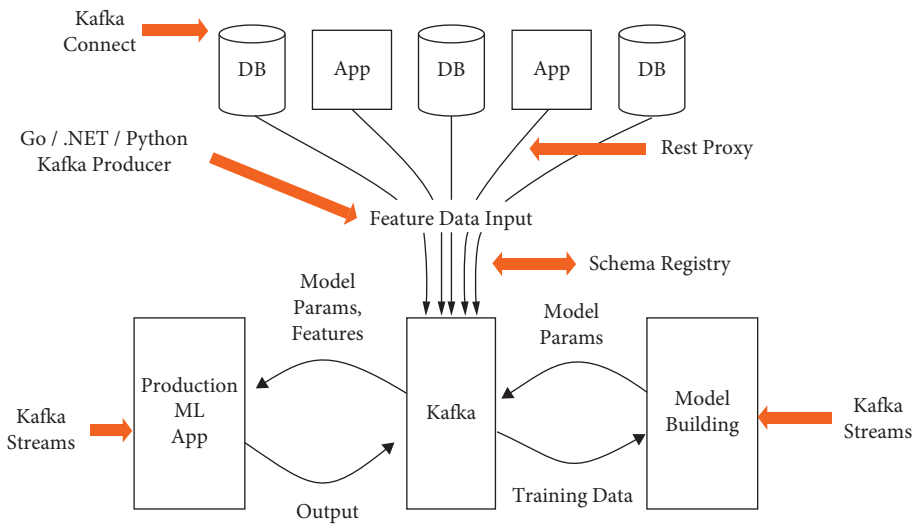


Figure 1: Multisource data structure.

orthogonal projection map, digital elevation model, and various digital thematic maps [6]. The integration and fusion of multivariate data are the main problem to be solved in the construction of spatial information system as shown in Figure 3 model values.

Multisource data generally refer to the diversity of data sources, for example, computer networks, cameras, questionnaires. At the same time, the heterogeneity of multisource data is the fundamental reason for the differences in data structures [7]. On this basis, the data structures are divided into three categories [8]. These three types have corresponding data models to ensure the consistency and coordination of data information and music emotion. These three types of data include the following:

(1) Data structure represented by e-government tabular data, and different information, such as name, occupation, income, is usually aggregated with the ID of the person or institution as the anchor point. Subsequently, a series of organizational forms such as basic library, theme library, and theme library will evolve.

(2) Unstructured data represented by video, image, voice, and text, most of them need to be analyzed and processed into structured data before they can be used.

(3) Spatiotemporal data represented by geographic information, IOT, and trajectory data.

*2.2. Practical Application of Multisource Data in Music Emotion.* Multisource data can interactively extract multiple types of music emotion data, which can not only improve the data value and use the value of music emotion, turn it into a visual data model, and be used in music emotion representation learning, which can help learners understand and learn music more intuitively and vividly [9]. Multisource data not only analyzes all kinds of information in music emotion but also combines big data technology to deeply mine the key information in music emotion, so as to find the special connections in music emotion. Even different types of music emotion can be deeply mined. The advantage of multisource data is also reflected in the efficiency of music emotional learning. Due to the integration of all kinds of data information, the final results can be obtained through accurate calculation. The results can guide the learning of music emotional representation and promote the sensitivity of emotional representation [10].

*2.3. To Characterize Learning Methods.* In other methods of representation learning, we usually consider representations that are easy to model, such as if all entities are sparse or independent of each other. First, we look at if semisupervised learning $p\ (x)\ p\ (x)$ does not benefit from learning $p\ (y\ x)\ p\ (y\ x)$. If the sample $p\ (x)\ p\ (x)$ is normally distributed, we want to learn $f\ (x) = E\ [y\ x]\ f\ (x) = E\ [y\ x]$. Obviously, just from the observation training set $x$, we cannot provide us any useful information. Next, looking after how the lower half of supervised learning becomes useful. If $x$ is mixed by different parts, different $y$ parts will constitute different $x$ values, as shown in the figure below. If the component y constituting x is completely independent, then modeling $p\ (\mathrm{x})\ p\ (\mathrm{x})$ accurately re-partition component, the single-label sample for each classification will be sufficient to learn$p\ (y\ x)\ p\ (y\ x)$, as shown in Figure 4.

An unsupervised manner without label samples has been able to obtain factor $y$. If $y$ is closely related to a certain promoter of $x$, then $p\ (y)\ p\ (y)$ and $p\ (y\ x)\ p\ (y\ x)$ are also strongly correlated, and unsupervised representation learning associated with variable latent factors will also be useful as a semisupervised learning strategy. Given our hypothesis is that $y$ is one of the $x$, now let $h$ represent all. The real generation process can be obtained through the following directed graph model, $h$ as the ancestor of $x$ (the relationship in the graph is $h\text{-- }>x$) $p\ (h, x) = p\ (x \mid h)\ p\ (h)\ p\ (h,\ x) = p\ (x \mid \mathrm{h})\ p\ (h)$. The result is that the data have a marginal probability $p\ (x) = \mathrm{Ehp}\ (x \mid h)\ p\ (x) = \mathrm{Ehp}\ (x \mid h)$.

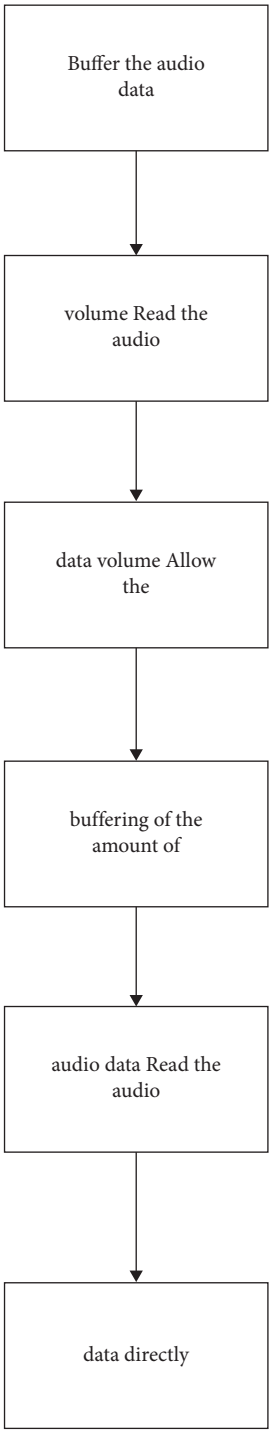


Figure 2: Buffer of the audio data volume.

From the above intuitive observations, we can conclude that the best model for expressing $x$ is the one that can reveal the true structure, where $h$ is the latent variable that explains the observed variables in $x$. The perfect representational learning discussed above should reproduce these latent factors. If $y$ is one of them (or is closely related to one of them), it will be easy to learn to predict $y$ from this statement [11, 12]. We can also see that the conditional distribution of a given $x$, $y$ is constrained by Bayesian laws (the above equation). Figure 5 shows the music emotion representation learning process.

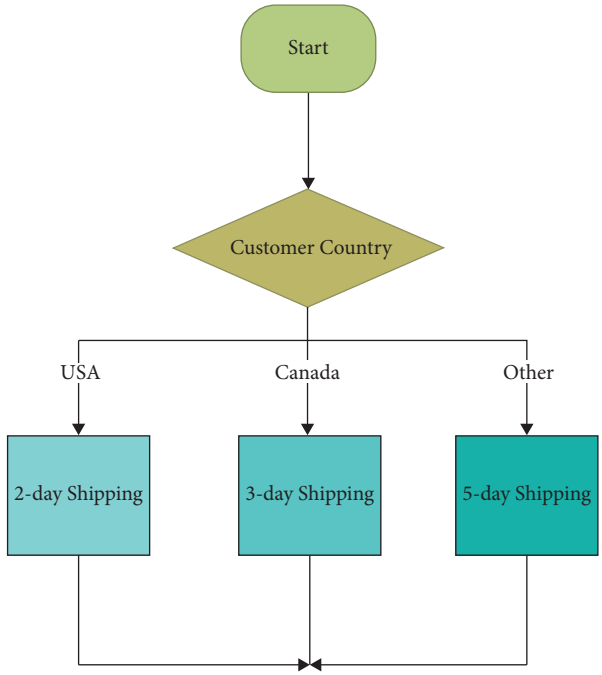


Figure 3: Information system model.

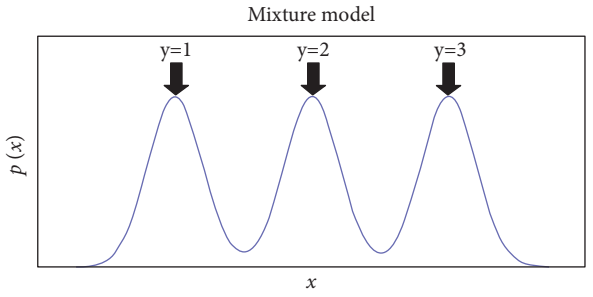


Figure 4: Mixture model.

## 3. Music Emotion Analysis

*3.1. Sentiment Analysis.* The precondition of music induction is full music emotion. Music emotion is the core of music. When music emotion reaches a certain extent, the music displayed will be deeply rooted in the hearts of the people and full of emotion, making the study of music emotion representation more perfect. Music can express the feelings of others more strongly. Music emotion has four obvious characteristics. First of all, music emotion can drive the brain's cognition and feeling and make each organ of the brain get a new experience and cognition, for example, dopaminergic system, cerebral cortex, sensory system

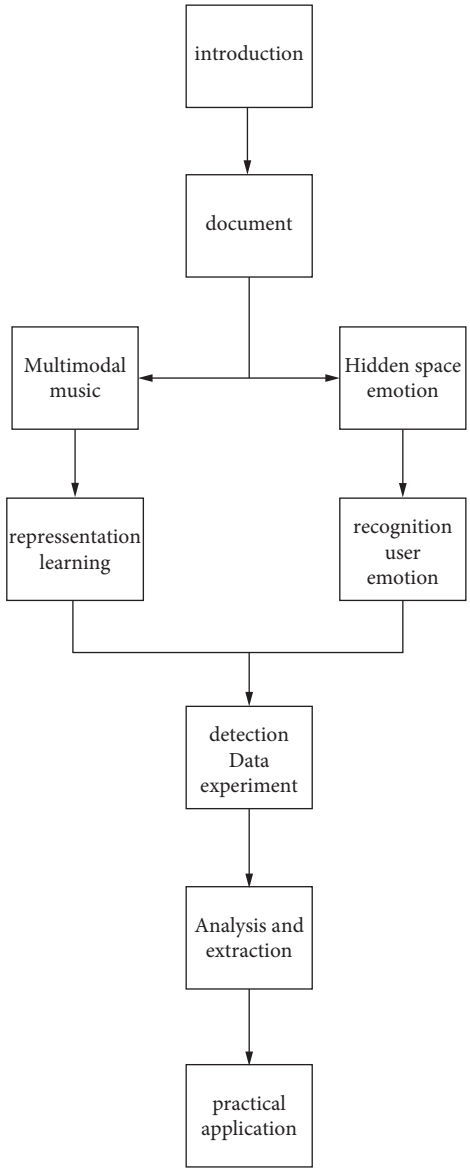


Figure 5: Music emotion representation learning process.

[13, 14]. The second is that we rarely experience negative emotions such as guilt, shame, jealousy, disgust, contempt, embarrassment, anger, and fear in music emotions. The emotional experience brought by music is generally considered to be positive and consistent in the world. Third, the emotional potency of music itself is not enough to determine whether people like a piece of music. Sad or sad music is also addictive. Figure 6 shows the emotional map.

*3.2. Emotional Analysis Algorithm.* Linear order is that with the expansion of the input scale, the corresponding calculation time increases in a straight line. The code in the following algorithm loop needs to be executed $n$ times, so the large $O$ meter method is $O(n)$.

```
private void calculate1() {
    int sum = 0;
    int n = 100;
    for (int i = 1; i ≤ n; i++) {
        sum+ = i;
    }
}
```

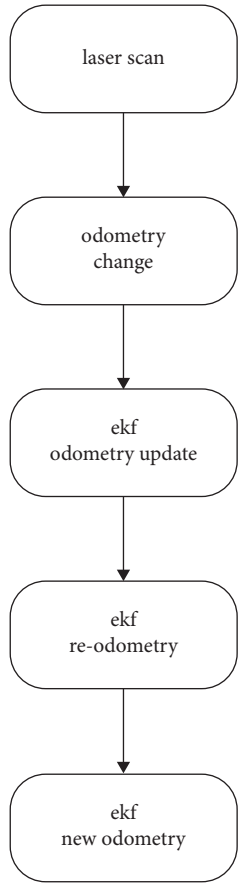


Figure 6: Emotional map.

*3.2.1. Code 1.* The $i$ cycle of the outer layer needs to execute $n$ times, and the $j$ cycle of the inner layer needs to execute $n$ times, so the large $O$ counting method is $O(n\hat{}2)$.

```
private void calculate2() {
    int sum = 0, n = 100;
    for (int i = 1; i ≤ n; i++) {
        for (int j = 1; j ≤ n; j++) {
            sum+ = i;
        }
    }
}
```

*3.2.2. Code 2.* The cubic order of the following algorithm in the outer layer $i$ cycle needs to execute $n$ times, the $j$ cycle of the inner layer needs to execute $n$ times, and the innermost layer $k$ cycle needs to execute $n$ times, so the large $O$ meter method is $O(n\hat{}3)$.

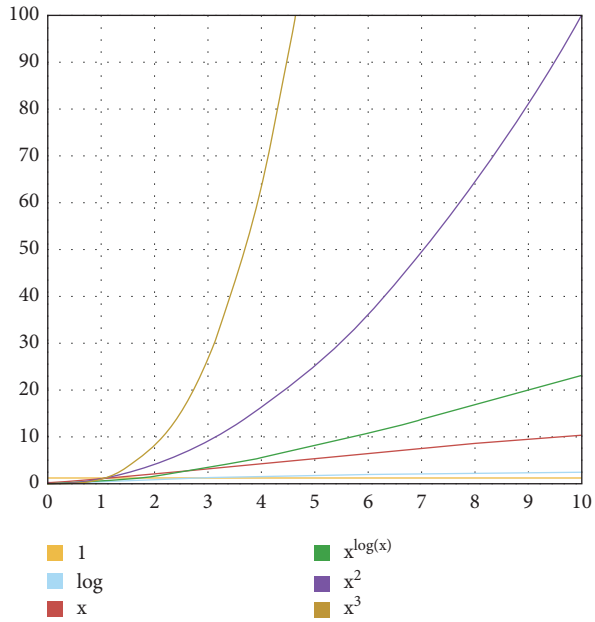


Figure 7: Simulation results.

```
private void calculate3() {
    int x = 0, n = 100;
    for (int i = 1; i ≤ n; i++) {
        for (int j = 1; j ≤ n; j++) {
            for (int k = 1; k ≤ n; k++) {
            }
        }
    }
}
```

*3.2.3. Code 3.* The logarithmic function log (*n*) can also be expressed as log2 (*n*), with large *O* is *O* (logn).

```
private void calculate4() {
    int i = 1, n = 100;
    while (i < n) {
        i = i * 2;
    }
}
```

*3.2.4. Code 4.* rOLinear order Linear order means that as the input scale expands, the corresponding computation time increases linearly. The code in the algorithm loop below needs to be executed n times, so the big-O method is O(n)

```
private void calculate5() {
    int n = 100;
    int i = n + 2;
}
```

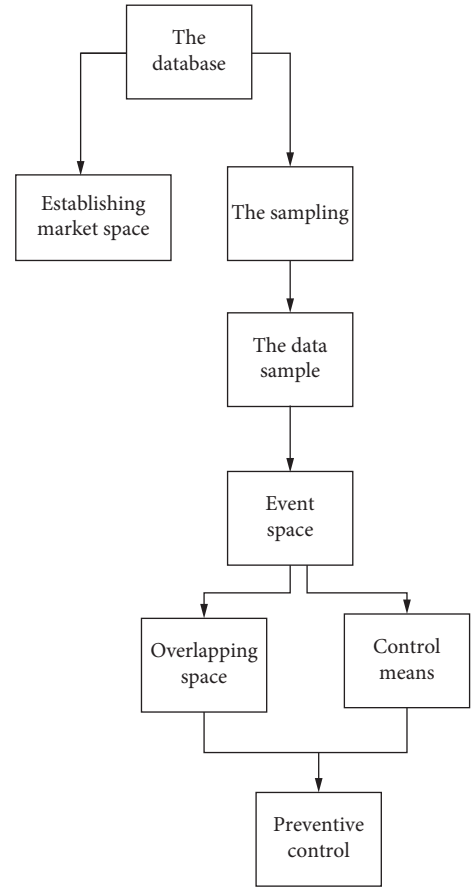


Figure 8: Feature extraction.

*3.2.5. Code 5.* The complexity of each *O* order is as follows: in summary and combined with the following function images [15], the complexity of various *O* order is as *O* (1) as shown in Figure 7.

The negative emotion generated in the music environment is "safe" [16]. No matter whether the emotion expressed by the music itself is happy or sad, the emotion induced by music can be happy. Fourth, musical emotion regulates almost all the activities of the limbic and paralimbic structures of the brain, including the hypothalamus, insula, and anterior cingulate cortex responsible for the awakening of the autonomic nervous system, the hippocampal region forming memory, and the prefrontal cortex involved in complex cognitive activities. Figure 8 shows feature extraction [17].

Music can not only reflect users' emotions and emotions but also use music to stimulate users' sense of identity, clearly convey the main emotions that music wants to show, attract users' attention, and relieve users' mood. In this process, the correlation analysis between multisource data and music emotion and music emotion further determines the guiding position of multisource data in music emotion representation learning [18]. At the same time, the active application of multimodal data makes the music emotion representation learning more concrete and visualized.

Representational learning is a process of transforming original data into data that is easier to be applied by machine

learning [19]. For input data, learning it to get new data or selecting the original data to get new data is called representational learning. The purpose of representation learning is to simplify the complex original data, refine the original data into better data expression, and make the subsequent tasks get twice the result with half the effort [20].

In addition to thermal coding, the representation learning of data information class also has distributed. Distributed representation is applied to text data for the first time, including the latent semantic index model (LSI), latent Dirichlet assignment model (LDA), and some variants. In recent years, with the emergence of deep learning, text-based distributed representation methods, namely word embedded representation model, have been proposed one after another. The distributed representation of text can not only greatly reduce the representation dimension but also contain the semantic information of the text; that is, the similarity between two words can be measured by the distance between word vectors [21].

Two models including word2vec are cbow and skipgram. Cbow predicts the words around music emotion, and skipgram predicts the words around according to the central word. Unlike word2vec, which only uses the local context window, gloschuh directly encodes the coexistence time of words and words into the coexistence matrix of word context and decomposes it through a specific weighted least square model to obtain an effective word vector representation. However, none of the above methods can correctly express polysemy, so Elmo suggests to solve the problem of polysemy expression by learning the context to express the text.

Network data are a kind of important data to represent relationships. Especially, in the era of big data, various complex networks and hundreds of millions of nodes and edges make the traditional representation based on adjacency matrix or adjacency table difficult to meet complex reasoning or prediction. Therefore, we use the low-dimensional vector to represent the nodes in the network and use it as the representation of the network for subsequent prediction tasks. This representation of the network is called network embedding. The traditional low dimensional vector representation method is called graph embedding, which generally adopts matrix decomposition to reduce the dimension, for example, by decomposing the Laplace matrix of the graph or the similarity matrix of the nodes. Figure 9 shows data flow to chart [22].

With the rise of deep learning, many existing methods apply the deep learning model or representational learning model to network representation to learn the representation of nodes. Deepwalk is the first network representation method based on deep learning technology [23]. It combines random walk and word embedding. It simulates the scene of text embedding by treating nodes as words and random walk paths as sentences, so as to use the existing text embedding model to learn network embedding. Inspired by deepwalk, line and node2vec appear. Line uses breadth first search strategy to generate context nodes, while node2vec directly extends deepwalk and adopts a preferred random walk process to combine depth first search and breadth first

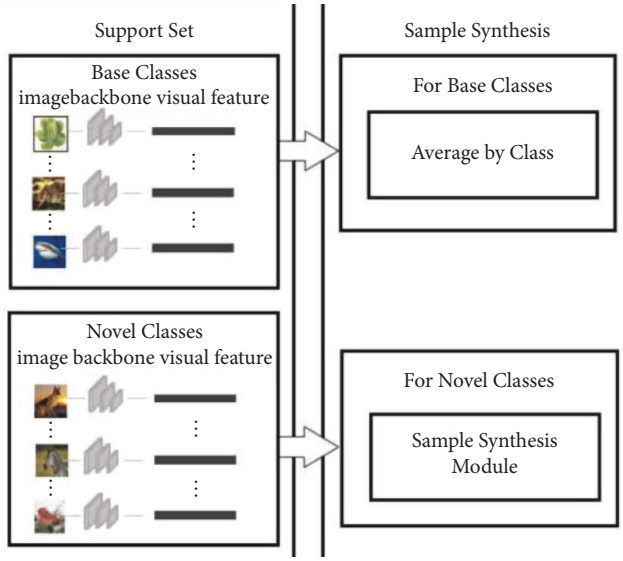


Figure 9: Data flow to chart.

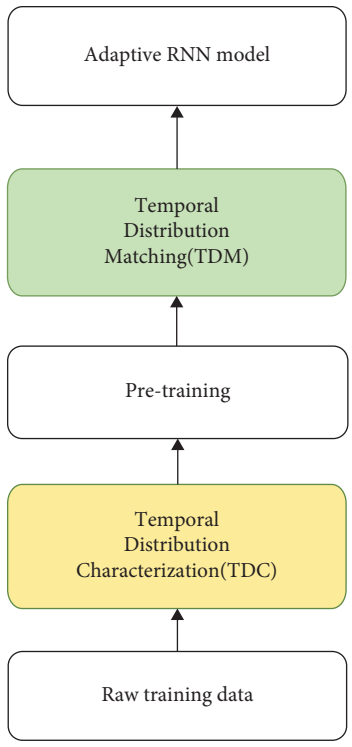


Figure 10: Multiple sequences.

search. In addition, sdne uses a deep automatic coder to preserve the local network structure; Dngr uses random surfing strategy to capture graphic structure information and then uses de manic automatic coder to learn node embedding and richer network representation principles. Figure 10 shows multiple sequences.

## 4. Music Representation Learning

*4.1. Music Hidden Space.* The first step of music representation learning is to find the hidden space of music [24]. The hidden space of music mainly refers to the nature that is not

Table 1: Music representation learning index.

| Total song list of the data set | Data value |
|---|---|
| Total number of songs in the dataset | 1621869 |
| Maximum number of songs on the playlist | 1000 |
| Minimum number of playlist | 1 |
| The average number of songs | 119.56 |
| Maximum number of labels on the playlist | 3 |
| Minimum number of labels | 1 |
| The average number of labels | 2.85 |

easy to express except music attribute, music emotion, music type, music content, music rhythm, etc., which is called hidden space. Hidden space is also the key part of hiding the internal relationship of music information [25, 26]. After determining the hidden space, multisource data and matrix decomposition are used to extract music emotion hierarchically. Finally, the hidden space information in music is expressed through accurate data algorithm, which fully proves the role of multisource data on music emotion [27] as shown in Table 1.

### 4.2. Music Data Analysis

(1) The core idea of music data analysis is to establish indicators + select analysis methods + output model results.

(2) Music data analysis can focus on four common perspectives: song correlation, leaderboard/starter list, singer correlation, and fan group.

(3) When performing song association analysis, the core index selected and constructed is "music style + singer + age + language + ordering other songs," and the analysis method is collaborative filtering.

(4) When analyzing the leaderboard/first list, the core index selected and constructed is "user ordered songs + leaderboard/first list songs + correlation between user ordered songs and leaderboard/first list." The analysis method is index threshold screening.

(5) When performing singer association analysis, the core index selected and constructed is "gender + style + age + region + ordering other singer songs," and the analysis method is association analysis or cluster analysis [28]. The calculation formula of music emotion recognition is as follows:

Data: playlists;$U$ 0, $V$ 0, $\lambda$ uI, $\lambda$vI, $\lambda$UC, $\lambda$vC, $T$.

Result: $U$, $V$.

(1) Construct matrix $G$ $u$ and $Gv$;
(2) Construct matrix $L$ $u = D$ $u - W$ $u$ and $Lv = Dv - Wv$;
(3) initialization $U = U$ 0, $V = V$ 0;
(4) while not convergent and $t \leq T$ do
(5) update matrix $U$ by using formula 3-thirteen;
(6) update matrix $V$ by using formula 3-nineteen;
(7) end
(8) norm $U$, $V$;
(9) return $U$, $V$;

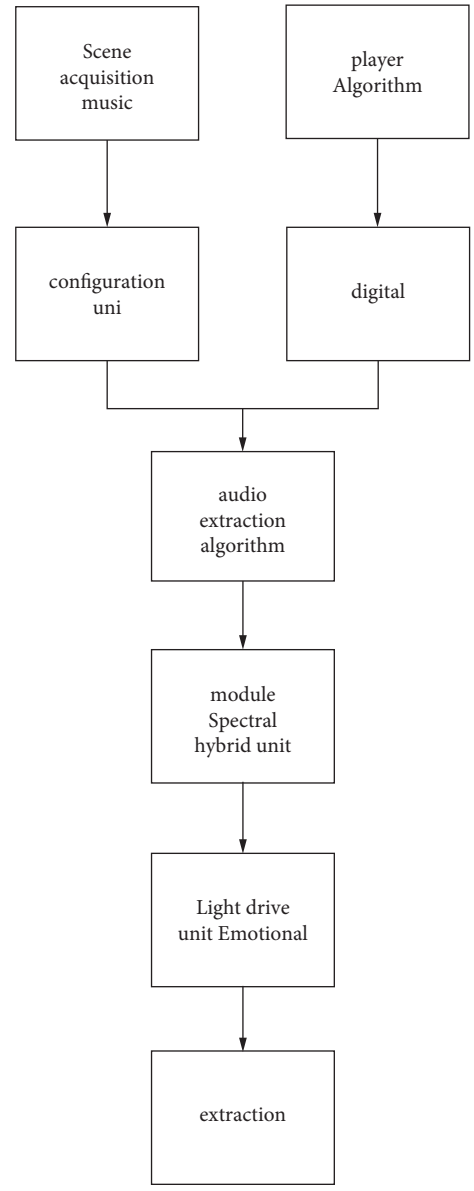


Figure 11: Multimodal.

*4.3. Music Emotion Mode.* Due to the continuous development of music emotion representation learning, digital music has attracted more and more attention [29]. The key features of music emotion are not only audio and lyrics but also the attributes of music itself. Especially with the promotion of the Internet, music data and music types are increasing. People enjoy music on different software platforms and save a lot of user data, which contains a lot of valuable information. However, most of this information comes from a wide range of sources and has complex structure [30]. It is necessary to classify and sort out this messy information. Therefore, we must use multisource data to manage this diversified data information and give full play to the role of multisource data. Figure 11 shows multimodal.

In order to effectively identify music emotion, a music emotion representation learning method based on user song listening list is proposed. Multimodal music performance

learning architecture, which mainly uses two technologies, integrates part of the information and multimodal model [31].

Analyzing music sentiment based on user listening song lists requires high workload and high labor cost. Functions such as media promotion, adding more users, listening to songs, creating song lists, and tagging favorite music information on the platform in music social software all generate a large amount of usable data related to music sentiment classification [32, 33]. The emotional perception of music has important research value. Therefore, this paper creates a user song list for the song list and deduces a decomposition algorithm based on the music emotion generation matrix from the marking behavior.

## 5. Conclusion

This paper aims to apply Internet multisource data to the study of emotional representations in music to achieve a more accurate study of user emotions. This paper combines the Internet data sources with the music emotion recognition technology and uses the Internet technology to extract and classify the music emotion content, which can better analyze the music emotion of learners [34]. Multisource data can effectively enhance music emotion recognition and use a variety of data mining algorithms to mine users' emotions, so as to quickly acquire users' music tendency to accelerate the application and in-depth development of the Internet in the music field. This paper not only verifies the effectiveness of Internet multisource data in the emotional representation of music learning but also deeply analyzes the problems existing in the emotional representation of music learning, and promotes the further development of Internet technology.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

[1] H. Zhu and Y. Wang, "Intelligent evaluation of structural safety performance based on multi-source data," *Building structure*, vol. 1-10, 2022.

[2] H. Wang, "Research on data fusion technology and its application in WSN," *Data communication*, vol. 2, pp. 18–22, 2022.

[3] X. Zhang and J. Liu, "Code representation based on deep learning and its application," *Computer science and exploration*, vol. 1-22, 2022.

[4] M. Zhang and C. Wei, "Temporal and spatial variation analysis of groundwater reserves in North China Plain Based on multi-source data," *Geodesy and geodynamics*, vol. 42, no. 5, pp. 505–509, 2022.

[5] J. Huang, Li Xin, C. Fang, Ru Cui, H. Li, and B. Du, "Research on a landslide terrain depth learning and recognition model based on multi-source data fusion," *Chinese Journal of geological hazards and prevention*, vol. 33, no. 2, pp. 33–41, 2022.

[6] Y. Sheng, Q. Guo, M. Liu, J. Lan, H. Zeng, and F. Wang, "User charging behavior analysis and charging facility planning practice based on multi-source data fusion," *Power system automation*, vol. 1-24, 2022.

[7] Y. Guo, G. Lei, L. Zhang, and Z. Lu, "Study on spatial distribution characteristics of plot ratio in Shenyang based on multi-source data," *Industrial Building*, vol. 1-11, 2022.

[8] L. Yan, "New ideas of music exchange and cultural communication -- on the international vision of Chinese music in the context of the Internet," *Media*, vol. 6, p. 97, 2022.

[9] K. Zhang, C. Hu, Yu Song, H. min, Y. Zhao, and W. Chu, "Construction of Chinese obstetric knowledge atlas based on multi-source data," *Journal of Zhengzhou University (SCIENCE EDITION*, vol. 1-7, 2022.

[10] weidian Wu, K. Zeng, W. Zhou, Li Peng, and W. Jin, "Depth learning method of bus passenger flow forecasting based on multi-source data and response surface optimization," *Journal of Jilin University (Engineering Edition)*, pp. 1–16, 2022.

[11] S. Wan and L. Zhu, "Musical analysis and emotional expression of the art song in this dark grave," *Art evaluation*, vol. 3, pp. 59–61, 2022.

[12] S. Yan, "Research on emotion cultivation strategy in dance education and Teaching," *Journal of Jiamusi vocational college*, vol. 38, no. 2, pp. 77–79, 2022.

[13] Q. Song, "Design of music recommendation system based on interest emotion algorithm," *Microcomputer application*, vol. 38, no. 1, pp. 82–84 + 88, 2022.

[14] W. Bo, "Hero images interpreted by different emotional intersection-music emotional appreciation of Aria thousands of miles of spring all over the home," *Voice of the Yellow River*, vol. 23, pp. 64–67, 2021.

[15] X. Xu, J. Zhang, and M. Nian, "Online game traffic recognition based on representation learning," *Computer system application*, vol. 30, no. 12, pp. 172–179, 2021.

[16] J. Kang, H. Wang, G. Su, and L. Liu, "A review of music emotion recognition," *Computer engineering and application*, vol. 58, no. 4, pp. 64–72, 2022.

[17] X. Zhu and B. Xu, "Emotional semantic matching algorithm of music and words based on Gru algorithm," *Computer technology and development*, vol. 31, no. 11, pp. 46–51, 2021.

[18] L. Chang, "Constructing a bridge to enrich inner music emotion -- the important position of music hearing training in Solfeggio and ear training course," *People's music*, vol. 8, pp. 65–67, 2021.

[19] J. Zhao, H. Liu, X. Liang, and Y. Gao, "Multimodal music emotion recognition based on the combination of knowledge distillation and transfer learning," *Journal of Fudan University (NATURAL SCIENCE EDITION)*, vol. 60, no. 03, pp. 309–314+322, 2021.

[20] N. Hui and N. Jiangping, "Music emotion appreciation classification based on feedforward neural network multi feature fusion algorithm," *Microcomputer applications*, vol. 37, no. 2, pp. 91–94, 2021.

[21] Y. Chen, K. Zhang, and L. Liu, "The "musical personality" of artificial intelligence music -- endowing the dynamics of music creation to enhance the emotion of music," *China National Expo*, vol. 3, pp. 130–132, 2021.

[22] X. Li, L. Han, and J. Li, "Zhou JingweiMultimodal music emotion classification based on optimized residual network," *J]Computer and modernization*, no. 12, pp. 83–89, 2020.

[23] Z. Wang, "Music emotion intelligent classification algorithm based on feature word location factor," *Electronic design engineering*, vol. 28, no. 17, pp. 56–60, 2020.
[24] J. Dong, "Emotional interaction, classroom application -- Research on the application of emotional teaching in junior middle school music classroom," *Intelligence*, vol. 18, p. 43, 2020.
[25] Lu Ni, "Design of music emotion classification method based on dual modes of audio and lyrics," *Automation technology and application*, vol. 39, no. 5, pp. 166–169, 2020.
[26] D. Zheng, "Multi feature fusion music classification algorithm based on deep confidence network," *Electronic design engineering*, vol. 28, no. 4, pp. 132–136, 2020.
[27] D. Yu, "Cultivation of emotion and musical aesthetic ability in Vocal Music Teaching," *Northern music*, vol. 2, pp. 137-138, 2020.
[28] D. Zheng, "Music emotion recognition and classification algorithm based on feedforward neural network," *Information & Technology*, vol. 43, no. 12, pp. 57–61, 2019.
[29] J. Liu and J. Liu, "Research on computer analysis and automatic recognition technology of music emotion," *Northern music*, vol. 38, no. 9, p. 220, 2018.
[30] C. Yu, "Research on transfer learning method in music emotion recognition," *Modern computer (Professional Edition)*, vol. 6, pp. 3–6, 2018.
[31] An Gong, M. Ding, and F. Dou, "Multi feature fusion music emotion classification method based on DBN," *Computer system application*, vol. 26, no. 9, pp. 158–164, 2017.
[32] X. Kong, "The importance of strength contrast to music emotion expression in piano performance," *Art evaluation*, vol. 17, pp. 47-48, 2017.
[33] X. Chen and Y. Deshun, "Research progress of music emotion recognition," *Journal of Fudan University (NATURAL SCIENCE EDITION)*, vol. 56, no. 2, pp. 136–148, 2017.
[34] Y. Wu, deran Liu, and X. Xu, "Design of real-time music recommendation system based on two-way emotion analysis," *Journal of Dalian University for nationalities*, vol. 19, no. 1, pp. 76–79, 2017.