Research Article

Chunhai Li and Xiaohui Zuo*

Classical music recommendation algorithm on art market audience expansion under deep learning

Abstract: The purpose of the study is to help users know about their favorite music and expand art market audiences. First, the personalized recommendation data of classical music are obtained based on the deep learning recommendation algorithm technology, artificial intelligence, and music playback software of users. Second, a systematic experiment is conducted on the improved recommendation algorithm, and a classical music dataset is established and used for model training and user testing. Then, the network model of the classical music recommendation algorithm is constructed through the typical convolutional neural network model, and the optimal parameters suitable for the model are found. The experimental results show that the optimal value of the dimension in the hidden layer is 192, and 24,000 training rounds can converge to the global optimum when the learning rate is 0.001. The personalized recommendation is provided for target users by calculating the similarity between user preference and potential features of classical music, relieving the auditory fatigue of art market audiences, improving user experience, and expanding the art market audience through the classical music recommendation system.

Keywords: deep learning, classical music, recommendation algorithm, the increase of the size of the audience

1 Introduction

Nowadays, the rapid development of the Internet brings a lot of information for people to choose from. However, it also causes the problem of information overload, resulting in great competitive pressure on information producers [1]. For the collector of information resources, it is difficult to pick out valuable and suitable information from a large deal of information. In this case, the time used for selecting information is greatly increased, reducing the working efficiency. Meantime, this leaves people drown in excessive and invalid information. For information producers, it is increasingly hard to enhance the competitiveness of their production of information, to stand out in the vast amounts of information, and to win the attention of the public [2]. The information overload becomes more and more severe. In view of this, the two complementary tools of search engines and recommendation systems come into being. Since the application of a search engine is passive for the user, it is required to know the personalized needs and characteristics of each user. The recommendation system automatically recommends the keywords and submits them to the search engine, showing users relevant results in the system after they are obtained [3].

* **Corresponding author: Xiaohui Zuo,** Institute of Education Sciences, Wuhan University, Wuhan, 430072, China,
e-mail: 305903169@qq.com
Chunhai Li: School of Art, Chongqing College of Humanities Science & Technology, Chongqing, 401524, China; Institute of Education Sciences, Wuhan University, Wuhan, 430072, China, e-mail: 191320@whu.edu.cn

With the continuous development and improvement of Internet technology, more and more classical music-related industries have begun to turn to online music platforms to publicize their music [4]. The appearance of platforms, such as Netease Cloud, Tencent, and KuGou, provides channels for increasing the size of the audience of classical music. The online listening, downloading, purchasing, sharing, and other functions of the music platforms greatly improve the speed and enlarge the scope of classical music, which is helpful in giving more users access to the resources of classical music. However, there are more and more pieces of music in the music library, which leads to more time and energy that the users have to spend finding their favorite music [5]. In terms of the search engine, the traditional search method usually retrieves the names of classical songs, singers, albums, and so on, which ignores the user's personalized differences and causes the "long tail" phenomenon of classical music. To solve this problem, the recommendation systems of classical music should be applied, because they can collect users' personalized information and preferences and analyze the data of classical music, recommending the classical music that conforms to users' preferences [6].

Based on convolutional neural network (CNN) and deep learning (DL), users' personalized preferences and the characteristics of classical music are collected, and the personalized recommendation system of classical music for each user after the data are analyzed. The basic principle of the recommendation system of classical music is as follows: (1) the audio resources of classical music in the system are processed, and then the spectrum and note characteristics of classical music are extracted; (2) the music is divided into several segments, and the probability distribution of each segment on the preset classification is used as a new classification basis; and (3) the traditional classification of classical music and user preference, combined with the data of the user's listening, collection, praise, sharing, purchase records, a recommendation system is designed to automatically identify the user preferences and characteristics. Data analysis on the application of this recommendation system shows that it plays a great role in increasing the size of the audience in the art market. Since audiences are one of the main driving forces for the development and progress of the art industry, the results have great significance and open up a bright prospect for the dissemination of classical music.

This article is divided into four sections. First, Section 1 shows the research background and the main content of the study; Section 2 introduces the research methods and data needed in this study; Section 3 conducts the comparative analysis of classical music recommendation algorithm and data; and Section 4 is the conclusion, which summarizes the main research work of this study and shows the direction of future research work.

2 Method

2.1 Recommended system of classical music

In the literature research on the music recommendation algorithm, Shi proposed a music recommendation method that combines the long-term, medium-term, and user's real-time behavior, considers the dynamic adjustment and the influence weight of the three behaviors, and uses advanced long-term and short-term memory technology to improve the effectiveness of the music recommendation [7]. Dharsini et al. built an efficient music recommendation system that uses facial recognition technology to confirm users' emotions [8]. Jin and Han proposed a music recommendation algorithm combining the clustering and potential factor models. First, users' music playback records are processed to form the user music matrix. Then, the probability model of potential factors is used to analyze the data on the result matrix, and the user preference matrix U and the music feature matrix V are obtained. On this basis, two clustering algorithms are used for user clustering and music clustering of the two matrices. Finally, the user-based collaborative filtering algorithm is used to predict users' preference matrix and commodity feature matrix that complete the clustering [9].

The recommended system of classical music mainly includes three parts: people, model, and results. Here, based on the actual needs, the recommendation model of classical music is divided into two categories, namely the user preference model and the music resource model. In terms of the user preference model, the recommended algorithm is based on context, which requires the use of music labels, user data, and result feedbacks to establish its model. In terms of the music resource model, the recommended algorithm is based on content, which needs to establish a model based on the characteristics of classical music, such as category, emotion, and melody. After the two models are established, collaborative filtering is applied to realize the personalized recommendation [10].

Deep learning (DL), an increasingly popular research method, is applied to the recommendation system of classical music for the improvement and optimization of algorithms, which was published in the top academic journal *Science* in 2006 and recognized by the academic community. This pioneering result based on DL establishes Hinton's authority in the field of DL. The study draws two conclusions: first, the multi-layer neural network can better identify data compared with the single-layer, especially in the learning and performance of data features. Second, the progressive nature of DL can greatly reduce the difficulty of neural network training, so that the gradient decline of the neural network becomes more stable [11]. Furthermore, in terms of the application of DL, Google published its research on the application of a recommendation system (Google Play) based on DL and the algorithm model for YouTube videos in 2016 [12].

2.2 Increasing the size of the audience

Since the concept of "increasing the size of audience" appeared in the 1980s, it has been widely used in music, drama, dance, film, television, and visual art. Although it has no unified definition, the connotation of increasing the size of the audience is understood from the definition and description by many scholars based on different art categories and their own experiences. Christian Watt, an English scholar, argued that increasing the size of the audience refers to a powerful process to improve the service for existing audiences and attract new audiences. It is not a simple behavior process but a planned and targeted management process involving all aspects of the operation of museums to achieve their overall purpose to a high standard [8]. According to Ebrahimi et al., increasing the size of the audience is defined as enriching the audience's experience, helping them learn more, and deepening their enjoyment of art museum services [13]. Yan thought that the goal of broadening audiences is to establish a relationship between man and art [14]. Du summed up the increasing size of the audience as "an art to win new audiences and retain old audiences" [15]. Liao deeply analyzed the news selection, decoding, and coding abilities that news audiences should have based on news consumption features and the corresponding problems [16].

Here, the object of the research on broadening the audience is the art market, which is wider, compared with art museums, theatres, and stages. The concept of audience is deeper and broader. Besides, the "audience" in this study refers to art viewers and appreciators in the narrow sense, as well as art followers, learners, lovers, consumers, and the target audience of the whole art industry.

2.3 Recommendation algorithm based on deep learning

2.3.1 Deep neural network model

DL is proposed under the upsurge of machine learning. It opens up a more convenient and broad perspective for research and becomes one of the most popular research methods. It differs from other machine learning because it has multiple learning methods, which can be divided into several ranks and successively arranged [17]. Each learning method is composed of some simple and nonlinear modules, and these modules can transform the previous module representation into a higher level of abstraction. In the complex combinational

transformation, the learning system can perform more complex tasks, like simulating the perception process of the human brain neural network on external stimuli. In terms of classification, a higher level of representation can magnify the irrelevant changes instead of controlling them when identifying important features [18].

The research on the deep neural network model is as follows: Mun et al. developed a deep neural network model, which can estimate and quantify gait spatio-temporal parameters from foot features [19]. Ren et al. proposed to use GA to automatically iteratively generate the most suitable network model based on existing datasets, remove redundant nodes and connections of the original network model, and make the optimized model more streamlined according to the fact that there are too many human experience factors involved in the training and model compression of DL networks [20]. Lee and Lee developed a process-centered assessment method using the concept of the deep neural network and a series of facial images [21].

The principle of the deep neural network is to establish a suitable learning network so that many hidden layers in the network can learn various features actively. Hidden layers have a progressive relationship with each other, which is an abstract description of the last layer. The analysis of data samples by the deep neural network is deep and more accurate, and the classification is more reasonable [22]. The deep neural network mainly includes three levels: the first is the input layer, which inputs the original data samples into the neural network training. The second is the hidden layer. After acquiring and analyzing the characteristics of the original data sample, the neural network compresses and abstracts the characteristics fully. The third is the output layer, which classifies the results of the analysis of the characteristics. The probability of each classification in the expectation is calculated [23]. The results are compared to the expected in the training, and the differences between the results are transmitted back.

2.3.2 Deep neural network model

In the application of common DL systems, the weights corresponding to each neuron represent their respective characteristics for learning in the network [24]. Linear models are usually used to memorize units. To fit external stimuli better, a nonlinear activation function can be added to the original linear model. In equation (1), the Sigmoid function is an activation function added to compress the continuous input to [0, 1] in the application

$$\sigma(z) = \frac{1}{1 + e^{-z}},$$
(1)

where z represents the linear model of neural units:

$$z = \sum_j w_j x_j + b.$$
(2)

CNN is a typical deep neural network. It combines the advantages of image processing and DL, improves the accuracy of feature recognition, and greatly reduces the computation of neural networks. It can be applied to image recognition and recommendation systems.

2.3.3 Feature extraction method of the notes of classical music

The classical music signal is composed of pitch and pantone. The fundamental pitch determines its pitch. Therefore, the detection of the pitch period is the key to the recognition of the notes of classical music [25]. The pitch detection method based on the autocorrelation function is a simple and classical time-domain detection algorithm. The definition of the short-term autocorrelation function is shown in equation (3)

$$R_i(k) = \sum_{m-1}^{n-m} y_i(m) y'_i(m + k),$$
(3)

where R_i is the covariance and y_i is the autocorrelation function.

The frameshift method is used to select a reasonable pitch period, reduce the interference of frequency doubling waves, improve the accuracy of note recognition, and add the note characteristics to the spectrum sample to obtain the note spectrum [26].

2.4 Comparison of three recommendation algorithms

Knees compare content-based and context-based recommendation algorithms, which are mentioned in the literature review. Here, the two algorithms are compared with a DL-based recommendation algorithm, as shown in Table 1.

Table 1: Comparison of the three recommendation algorithms

	Prerequisite	Metadata	Cold start problems	Preference	Characteristics
Content-based	Music files	Need	N.	N.	Objective, direct, and numerical
Context-based	Users	Do not need	Y.	Y.	Subjective, noisy, semantic
The algorithm based on DL	Music files	Need	N.	N.	Objective, fuzzy, and extensible

DL-based data processing is the same as content-based data processing, extracting features from audio metadata. Although it avoids the cold start problem, the method increases the burden on data processing [27]. This is because the capacity of a piece of music is far greater than the amount of information generated by users. In this case, context-based recommendation algorithms like collaborative filtering are popular in the industry, and they are not only easy to handle and deploy, and independent of metadata, and can explore the preferences of users. However, the cold start is always a problem in the context-based recommendation algorithm, which is not applicable if there is little information available [28].

In addition, the recommendations are more likely to be similar to their preferences rather than related. The context-based recommendation can only provide the user with the preferences of other users, as some business platforms do. For example, people who have listened to the song will also like the following ones. This recommendation is not an accurate assessment of the similarity between the two songs [29].

Extracting music features from audio signals can essentially reflect the types of music, which is more in line with the intuitive feelings of human beings on music. Content-based and DL-based recommendation algorithms have this advantage. Also, DL-based recommendation algorithms can apply to a broader platform because of the strong scalability of the DL model [30].

Based on the above analysis, it is concluded that each algorithm has inevitable defects. Thus, it is suggested to combine the recommendation algorithms. However, the combined algorithm is very complex. After the three recommendation algorithms are re-examined, it is found that content-based can collect user behavior data for the recommendation.

2.5 Artificial intelligence (AI)

As a technology, AI mainly studies the characteristics and laws contained in human intelligence activities. Based on these characteristics and laws, it imitates and constructs an artificial system with a certain degree of intelligence and attempts to make computers complete the tasks that require human intelligence. In short, AI analyzes how to use computer hardware and software to simulate the basic theory, method, and technology of human intelligent behavior through intelligent algorithms, platforms, or machines to simulate and extend human intelligence.

Machine learning can mine the effective data and association in large amounts of data through the network model so that the program has the ability of self-learning and self-prediction. As an algorithm for machine learning, DL can continuously optimize the structure of its network model by simulating the neural network of the human brain and extracting more high-quality data and connections. Machine learning and DL greatly promote the realization and development of AI in many fields, such as data mining, natural language, and computer vision.

2.5.1 Machine learning

Machine learning is the core of the basic technical level of AI. It mainly studies some behaviors of computer simulation human learning to absorb new knowledge and skills. Machine learning is derived from the early research field of AI. Learning methods can be divided into supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. Machine learning needs to continuously train the machine through a large amount of data. With the help of cloud computing technology, the machine manipulates the data to judge the results and help make decisions. Through data description, the machine can describe, identify, classify, and explain things or phenomena, helping human beings complete tasks.

2.5.2 DL

DL is an algorithm that can effectively realize machine learning. This algorithm is mainly used to simulate the neural network of human brains for learning and establish a network imitating the human brain mechanism to analyze data. Machines' analytical learning ability similar to human beings is finally achieved by learning the inherent laws and representation levels of sample data. The DL algorithm imitates the multi-layer neural network of the human brain, which can make the machine learn more complex feature rules from the low level to the high level, solving some complex and difficult problems more pertinently.

3 Research model and recommendation algorithm

3.1 Self-attention model

The self-attention model is a special attention model and is widely used because of its powerful feature extraction ability for sequence information. In the attention model, the three elements of Query, Key, and Value are different. However, in the self-attention model, Query, Key, and Value are the same, which are all equal to the input signal X. It can be understood that the essence of this operation is that each signal itself queries its importance in the group of signals or its association with other signals, and then the weight of the signal that is only related to itself is accumulated to obtain its self-attention value. Attention in machine translation can be intuitively reflected. Whenever a word is translated, the weight of Attention will be more on the input words associated with the word. If the self-attention model is transformed, that is, the source language and the target language are the same, what self-attention learns in DL can be observed intuitively.

The surface of the figure is the input sentence of self-attention, and the above is the output sentence obtained by self-attention calculation. It can be found that the input and output are the same, that is, it translates itself. The line segment in the figure indicates the weight of the calculated in the input sentence when the word is translated. The deeper the color is, the stronger the relevance is.

As mentioned above, the self-attention model can effectively model the pre-correlation and post-correlation of internal signals in a sequence and even learn the semantic relationship within a sentence. Since the acquisition of these association features by self-attention is not affected by the distance, indiscriminate association modeling can be carried out no matter how far the two signals are located.

3.2 Content-based recommendation algorithm

3.2.1 Content representation

A feature representation is calculated for each item, which is the core of the whole recommendation algorithm and shows the difference applied to different recommendation fields. It usually uses some attributes of the item itself to represent the item. For example, attributes such as authors and publishing time are easy to obtain, and the real representativeness is the specific content of the article itself. At this time, the specific content of the article needs to be transformed into structured data. The article is composed of words, and the weight information of each word can be used to form a vector. Article d can be expressed as follows:

$$d = \{w_1, w_2,...,w_n\}, \tag{4}$$

where w_i represents the weight of the ith word in this article, n is the size of the entire vocabulary, and the specific size of the weight can be obtained by term frequency-inverse document frequency. w_i can be defined as follows:

$$w_i = \frac{n_i}{\sum_{j=1}^{n} n_j} \log \frac{m}{m_i + 1}, \tag{5}$$

where n_i represents the number of times the ith word appears in article d, m represents the total number of articles, m_i represents the number of documents that appear in the ith word, and 1 is added to avoid zeroing errors. This shows that the algorithm believes that the more times a word appears in article d, and the fewer articles the word appears, the greater the weight of the word is, and the more representative the article d is, thus obtaining the vector representation of the article d.

3.2.2 Learning the features

The user's preference for new items is determined according to the user's favorite items. In the first step, the feature vector of each item has been obtained, and the user knows which items he likes and which items he does not like. At this time, a supervised machine learning binary classification task can be constructed to train a user model and then determine whether the user has a preference for new items. Here, the commonly used classification algorithms are decision trees, naive Bayesian, and neural networks.

3.2.3 Generation of the recommendation list

The optimal K-recommended items are obtained through the characteristics of the items obtained in the first two steps and the output of the user model. If the user classification model is used in the second step, only the items most likely to be liked by K users are obtained, or the user vector is obtained by averaging the feature vectors of all users' favorite items, and then the K items most similar to the user vector are obtained.

Since the content-based recommendation algorithm mainly analyzes the content attributes of items themselves, it can well solve the cold start problem of items. When a new item is added, it can still be recommended well, even if no user likes it.

4 Results and discussion

4.1 Experimental environment

The experiment is conducted on the servers that have Windows 10 operating systems installed. The server configuration is shown in Table 2.

**Table 2:** Hardware environment and parameters

Hardware environment	Processor	Mainboard	Memory	Graphics card	Basic frequency
Parameters	Intel i5-10400F	Asus B460 TUF Gaming	Weigang DDR4 3,000 Hz 16G	Seven Rainbow RTX3060 12G	2.20 GHz

Table 2 shows the hardware used in the study of the classical music recommendation algorithm. Python 3.6 is the programming language used to write the experimental code. The Tensorflow 1.12.0 DL framework is used to achieve the rapid construction of the model to improve the efficiency of code writing.

4.2 Dataset collection and data pre-processing

4.2.1 Data collection

In this study, Last.fm Dataset-1k user music dataset (http://www.last.fm) is used to train and test the proposed model. This dataset records the user's music listening behavior completely and is widely used in the research of the music recommendation algorithm, which is convenient to compare with other algorithms. The dataset collects 579,195 records from 300 users in the tsv format, and each record is presented in the form of six ancestors <user, timestamp, artid, artname, traid, traname>. The specific scale of the dataset is shown in Figure 1.
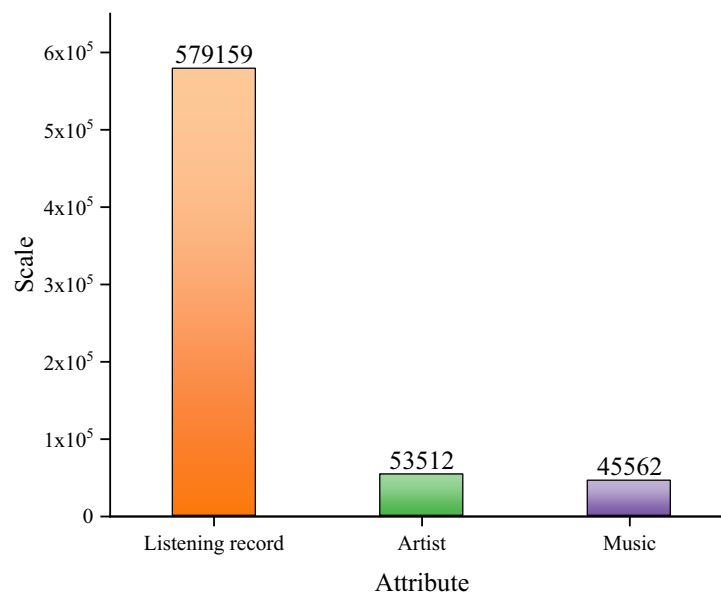
Figure 1: Last.fm Dataset-1k users datasets.

The following shows a record in the dataset with a specific example. User_000001 is shown in Table 3.

Table 3 shows that user_000001 listens to the song "Balloons of Love" sung by Jay Chow on 31 October 2020 at 3.41 p.m. 13 s, which also includes singer ID, avoiding the conflict between singers and songs.

Table 3: User_000001 the listening record

Attribute	User ID	Listening period	Singer ID	Artist	Song ID	Names of songs
Example	User_000001	2020-10-31T15:41:13Z	87c5dedd-371d-4a53-9f7f-80522fb73cb	Jay Chou	268b6266-29ce-4822-9f58-70034a8edb4a	Balloons of Love

4.2.2 Data pre-processing

Since the user's listening behavior has the characteristics of sessions, the user's listening records are divided into session records according to certain rules. Generally speaking, if the user does not listen for 40 min, his short-term preference may change. If the difference in the playing time of the two songs is 40 min, the former and latter records are divided into two session records based on this boundary. On this basis, the data need to be processed as follows, as shown in Table 4, to obtain a better recommendation effect.

After data are pre-processed according to the above rules, 208,627 session records are obtained. The average number of session records owned by users in the listening history is about 220.2, and the average number of each music session record contains about 8.4 songs, which meets the training data requirements of the model.

Table 4: Data pre-processing process

Operational procedure	Content
Remove the songs replayed	Users often play a song inadvertently and open a single loop play, and this is not the user's real intention. Even if it is the user's real intention, recommending the same song is meaningless, which will lose the diversity of recommendations and occupy valuable recommendation space. Therefore, the repeated songs are usually combined into one.
Remove short and long session records	Too short music session records have little significance for the training of the model and will cause too long _ PAD complement symbols, affecting the efficiency of model training. Because users often forget to turn off the player and play a lot of songs at random, users generally do not listen to a song for a long time, and too-long session records are more likely to be noisy and interfere with user preferences modeling. Therefore, session records between 5 and 40 are retained. Moreover, the model trained in such data can be good for effective recommendation even if encountering very short or very long sessions.

4.3 Hyperparameter setting

The architecture design of the classical music recommendation engine based on the CNN model and the DL model is shown in Figure 2.

The architecture of the recommended engine in Figure 2 contains four attributes:
1. Use user and databases to obtain user behavior features and song attributes.
2. User and music features are fed into the basic music recommendation algorithm to train the initial recommendation sequence after they are sorted.
3. Remove the initial recommendation sequence, obtain the features of the corresponding user and song, and splice the user listening song sequence features into the CNN model and DL model to train the TOP N songs to form the recommended list.
4. After the recommendation list is gained, it needs to observe the user's request for the recommendation list. When a personalized recommendation list is requested, the personalized recommendation list with new songs will be recommended to users. When users request popular songs, the list of popular songs will be recommended to users.
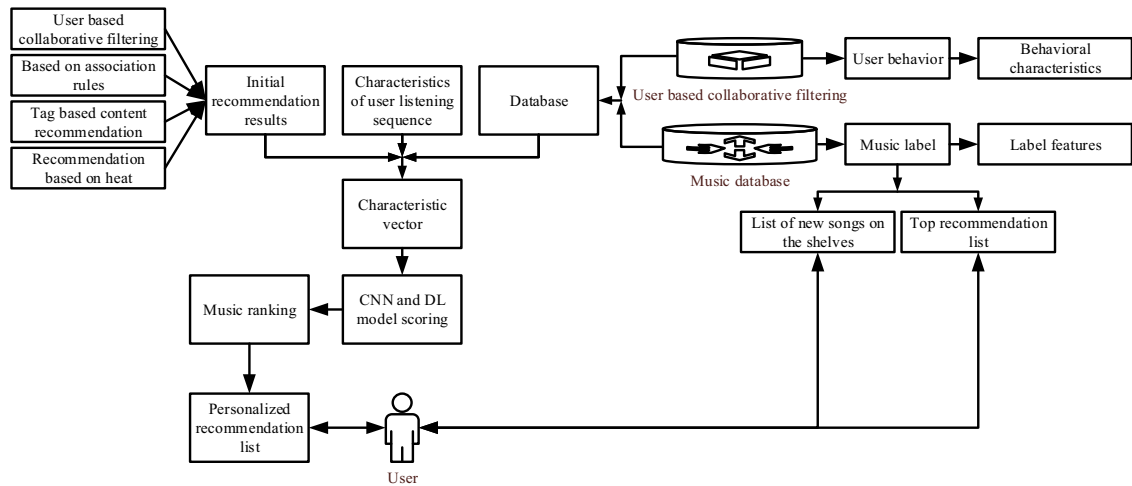
Figure 2: Architecture of the recommended engine.

4.4 Experimental results and performance evaluation

4.4.1 Comparative analysis of different dimensions in the hidden layer

The experiment aims to explore the influence of different dimensions in the hidden layer on the model and find the optimal value of the dimensions in the hidden layer. The number of dimensions in the hidden layer is often related to the feature extraction ability of the model. If the number of neurons in the hidden layer is too small, it will not extract enough information. If the number of neurons in the hidden layer is too large, it will introduce unnecessary noise to cause over-fitting and make the model more bloated. In this experiment, the dimensions in the hidden layer are set to 48, 96, 192, and 384, respectively. The experimental results are shown in Figure 3.
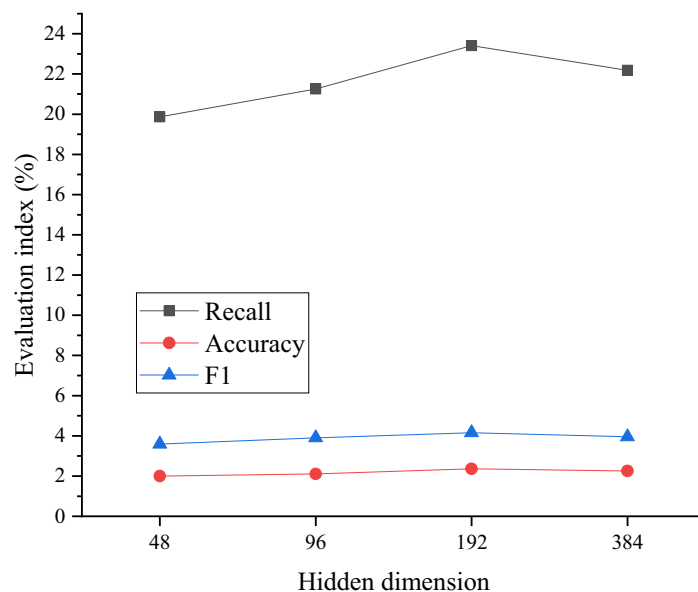

Figure 3: The changing state of index value under different dimensions in the hidden layer.

Figure 3 shows the influence of dimension in the hidden layer on the algorithm. When the number of the dimensions in the hidden layer is small, the model cannot carry or mine enough songs and user's preference information and cannot accurately predict the next song to be listened to by the user. When the number of dimensions in the hidden layer increases to 384, the overall effect of the model begins to decline, indicating that too many nodes in the hidden layer lead to the sparseness of users' preferences, producing much noise information and hurting the final result. Therefore, the optimal number of dimensions in the hidden layer is 192.

4.4.2 Comparative experimental analysis of different learning rates

The experiment is to explore the influence of different learning rates on the model. The learning rate determines the learning range of each training. A good learning rate can make the model converge to the optimal solution quickly and stably, while a bad learning rate may take a long time to converge to the optimal solution and even cause the model to be unable to converge. If the learning rate is too large, the model will oscillate at the lowest point and cannot reach the global optimum. If the learning rate is too small, it will lead to slow convergence or fall into the local optimum. When the optimal learning rate is searched for, the learning rate is set to 0.1, 0.01, 0.001, and 0.0001, respectively. The optimal learning rate can be found more quickly by scaling the multiple of 10, which is a common way to adjust the learning rate. The experimental results are shown in Figure 4.
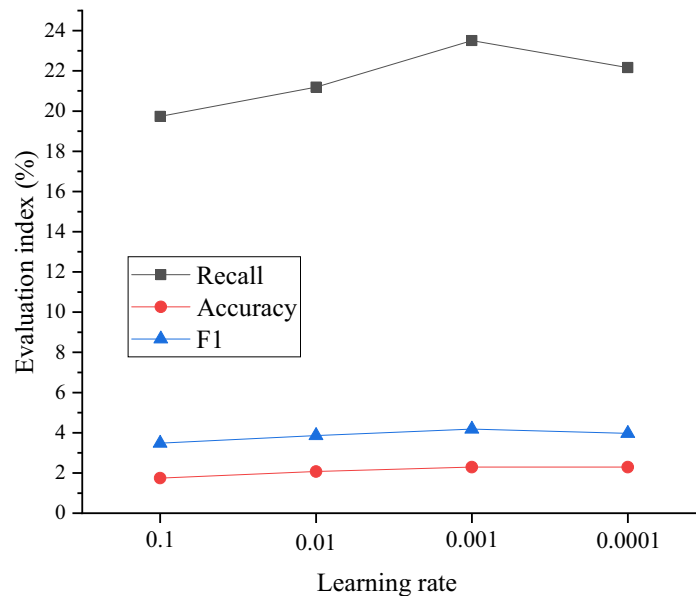

Figure 4: The changing state of indexes under different learning rates.

Figure 4 shows that the final effect shows a trend of first getting better and then getting worse as the learning rate decreases. When the learning rate is set to 0.1, the model oscillates at the lowest point, failing to converge to the global optimum. When the learning rate is 0.0001, the model does not converge to the global optimum. When it is close to the optimal point, it falls into the local optimum. At this time, 32,000 rounds of model training are needed to converge. When the learning rate is 0.001, 24,000 rounds of training can converge to the global optimum.

5 Conclusion

Based on DL, the recommendation algorithm is used to construct the classical music recommendation system. With the combination of DL and recommendation algorithms, various types of classical music are automatically extracted with the help of the relevant advantages of AI. In the audio of classical music, a higher level of feature representation can be obtained. Furthermore, the implicit characteristics of classical music are extracted to obtain user preferences by combining with the idea of the deep neural network of the classical music recommendation algorithm, as well as AI. Finally, the corresponding recommendation to each user's personalized preferences is obtained, achieving the purpose of increasing the size of the audience in the art market. According to the data of the recommended and shared classical music on music platform A, it can be concluded that the recommendation system of classical music can broaden the number of the audience in the art market, save the user's search time, and bring convenience to the user.

The shortcomings of the study are as follows: First, the recommendation system of classical music is only based on CNN, which may have some limitations. Second, the data collected have a few features, which are not comprehensive. Therefore, it is hoped that the design and establishment of the model can be further carried out by integrating other DL network models. Besides, the visualization method will be used to observe the learning state of each layer, improve the structure, and adjust the parameters, which can greatly improve the accuracy of the recommendation system. In terms of user characteristics, the issue time of classical music and user age will be added to the later research, making their characteristics more comprehensive.

References

[1]   Elbir AM. DeepMUSIC: Multiple signal classification via deep learning. IEEE Sens Lett. 2020;4(4):1–4.

[2]   Martin-Gutierrez D, Hernandez Penaloza G, Belmonte-Hernandez A, Alvarez Garcia F. A multimodal end-to-end deep learning architecture for music popularity prediction. IEEE Access. 2020;34(99):1.

[3]   Wen X. Using deep learning approach and IoT architecture to build the intelligent music recommendation system. Soft Comput. 2020;23(1):1–10.

[4]   Pandeya YR, Lee J. Deep learning-based late fusion of multimodal information for emotion classification of music video. Multimed Tools Appl. 2021;80(38):1–19.

[5]   Prisco RD, Zaccagnino G, Zaccagnino R. EvoComposer: An evolutionary algorithm for 4-voice music compositions. Evolut Comput. 2019;28(2):1–42.

[6]   Deshmukh P, Kale G. Music and movie recommendation system. Int J Eng Trends Technol. 2018;61(3):178–81.

[7]   Shi J. Music recommendation algorithm based on multidimensional time-series model analysis. Complexity. 2021;2021(1):1–11.

[8]   Dharsini SV, Balaji B, Hari K. Music recommendation system based on facial emotion recognition. J Comput Theor Nanosci. 2020;17(4):1662–5.

[9]   Jin Y, Han C. A music recommendation algorithm based on clustering and latent factor model. MATEC Web Conf. 2020;309(9):3009.

[10]   Schedl M. Deep learning in music recommendation systems. Front Appl Math Stat. 2019;5:44.

[11] Edwards JR, Borgstedt S, Barth B. New music recommendation algorithm facilitates audio branding. Mark Rev St Gallen. 2019;4:888–94.

[12] Pacha A, Haji J, Calvo-Zaragoza J. A baseline for general music object detection with deep learning. Appl Sci. 2018;8(9):1488.

[13] Ebrahimi AA, Abutalebi HR, Karimi M. A generalised two stage cumulants-based MUSIC algorithm for passive mixed sources localisation. IET Signal Process. 2019;13(4):409–14.

[14] Yan F. Music recognition algorithm based on T-S cognitive neural network. Transl Neurosci. 2019;10:123–34.

[15] Du X. Application of deep learning and artificial intelligence algorithm in multimedia music teaching. J Intell Fuzzy Syst. 2020;38(2):1–11.

[16] Liao BY. Composition and improvement strategies of news audience's media literacy in the omnimedia era. Contemp Soc Sci. 2020;24(4):128–37.

[17] Dorochowicz A, Kurowski A. Employing subjective tests and deep learning for discovering the relationship between personality types and preferred music genres. Electronics. 2020;9(12):2016.

[18] Oramas S, Barbieri F, Nieto O, Serra X. Multimodal deep learning for music genre classification. Trans Int Soc Music Inf Retr. 2018;1(1):4–21.

[19] Mun KR, Song G, Chun S, Kim J. Gait estimation from anatomical foot parameters measured by a foot feature measurement system using a deep neural network model. Sci Rep. 2018;8(1):9879.

[20] Ren HS, Bo XC, Ying XM. A deep neural network model compression method of diffuse large B cell lymphoma recognition based on genetic algorithm. Mil Med. 2018;42(10):757–61.

[21] Lee HJ, Lee D. Study of process-focused assessment using an algorithm for facial expression recognition based on a deep neural network model. Electronics. 2020;10(1):54.

[22] Huang Z, Jia X, Guo Y. State-of-the-art model for music object recognition with deep learning. Appl Sci. 2019;9(13):2645.

[23] Chowdhuri S. PhonoNet: Multi-stage deep learning for raga preservation in hindustani classical music. J Acoust Soc Am. 2019;146(4):2947.

[24] Briot JP, Pachet F. Music generation by deep learning - Challenges and directions. Neural Comput Appl. 2020;32(2):194–212.

[25] Gui R, Chen T, Nie H. The impact of emotional music on active ROI in patients with depression based on deep learning: A task-state fMRI study. Comput Intell Neurosci. 2019;2019(6):1–14.

[26] Purwins H, Li B, Virtanen T, Schluter J, Chang SY, Sainath T. Deep learning for audio signal processing. IEEE J Sel Top Signal Process. 2019;21:1.

[27] Sotiropoulos DN, Tsihrintzis GA. Artificial immune system-based music recommendation. Intell Decis Technol. 2018;14:1–17.

[28] Li T. Selection of audio materials in college music education courses based on hybrid recommendation algorithm and big data. J Phys Conf Ser. 2021;1774(1):012019.

[29] Mandloi K, Mittal A. Hybrid music recommendation system using content-based filtering and k-mean clustering algorithm. Int J Comput Sci Eng. 2018;6(7):1498–501.

[30] Gong W, Yu Q. A deep music recommendation method based on human motion analysis. IEEE Access. 2021;36(99):1.