Selection of Audio Learning Resources Based on Big Data

https://doi.org/10.3991/ijet.v17i06.30013

Peng Wang[1], Xia Wang[1](✉), Xia Liu[2]
[1] College of Music, Cangzhou Normal University, Cangzhou, China
[2] College of Computer Science and Engineering, Cangzhou Normal University, Cangzhou, China
wx19720513@126.com

Abstract—Currently, audio learning resources account for a large proportion of the total online learning resources. Designing and implementing a method for optimizing and selecting audio learning resources based on big data of education will be of great significance to the recommendation of learning resources. Therefore, this paper studies a method for selecting audio learning resources based on the big data of education, with music learning as an example. First, the audio signals were converted into mel spectrograms, and accordingly, the mel-frequency cepstral coefficient features of audio learning resources were obtained. Then, on the basis of the conventional content-based audio recommendation algorithm, the established interest degree vector of target students with respect to music learning was expanded, and a collaborative filtering hybrid algorithm for audio learning resources that incorporates the interest degrees of neighbouring students was proposed, which effectively improved the accuracy and stability in the prediction of students' interest in music learning. Finally, the experimental results verified the feasibility and prediction accuracy of the proposed algorithm.

Keywords—big data, audio selection, learning resources, interest degree

1 Introduction

In terms of learning, each student has a different level of cognition, motivation and learning habits. Luckily, intelligent tutoring systems, featured with high efficiency and low cost, provide learners with personalized potential learning space [1-11]. However, with the amount of online learning resources continuously growing, learners find it increasingly difficult to select the suitable learning resources from the massive and complex learning resource data and have to rely on the selection and recommendation from the learning management system of the online learning platform [12-17]. Currently, audio learning resources account for a large proportion of the total online learning resources, especially for students majoring in languages and music [18-22]. Therefore, designing and implementing a method for selecting audio learning resources based on the big data of education will be of great significance to the research on recommendation of learning resources.

Zhang and Kong [23] emphasized that there are a number of audio recommendation methods, and that different audio applications are used for different users according to different product features. On the platform, various recommendation methods need to be used flexibly. In music education in colleges and universities, in order to avoid improper selection of audio materials that may lead to the low enthusiasm of students in music learning, Li [24] adopted a hybrid recommendation algorithm incorporating big data - a personalized recommendation algorithm based on collaborative filtering recommendation algorithm (CF), obtained the user evaluation matrix based on big data, used the Pearson correlation coefficient to calculate the similarity between users and form the nearest neighbour set, and then obtained the nearest neighbour set of the target user, and generated the user-based recommendation set. Bogdanov et al. [25] presented three methods for music recommendation, two of which are based on semantic music similarity, and one based on the semantic probability model. This study attempted to achieve the visualization of users' music preferences by creating a human-like cartoon character, and preliminarily evaluated the technologies proposed by 12 subjects in the context of these applications. Deng and Leung [26] used a multi-dimensional emotional representation called resonance arousal valence to express musical emotions, and used an inverse exponential function to represent the emotion decay process, so that the relationship between acoustic features and affective reflexes based on this representation was well established. Bozzon et al. [27] proposes a new content-based music recommendation method, whose novelty lies in a similarity function, which does not consider the whole music pieces or their thumbnail representations, but instead analyzes the audio similarity between the semantic fragments from different tracks.

This paper summarized the common problems in the selection of resources by the learning management system of the online learning platform from the perspective of technical engineering implementation. Existing research has rarely considered data sparsity, cold boot and data noise. What is more, in real scenarios, a resource selection system has to compute a huge amount of data. How to efficiently optimize resources is currently an urgent problem to be solved. To this end, this paper studied the method for selecting audio learning resources based on the big data of education, with music learning as an example. The whole paper is organized as follows: (1) the audio signals were converted into mel spectrograms, and then the mel-frequency cepstral coefficient features of audio learning resources were obtained; (2) on the basis of the conventional content-based audio recommendation algorithm, the existing interest degree vector of the target students with respect to music learning was expanded; (3) a collaborative filtering hybrid algorithm for audio learning resources incorporating the interest degree of the neighbouring students was proposed, which further improved the accuracy and stability in the prediction of students' interest in music learning. The experimental results verified the feasibility and prediction accuracy of the proposed algorithm.

2 Feature extraction of audio learning resources

Different from information such as students' learning needs, behaviours and habits, audio learning resources cannot be directly selected by the recommendation system; instead, their features must be extracted first. In this paper, the audio signals were converted into mel spectrograms, and then the mel-frequency cepstral coefficient features of audio learning resources were obtained. Figure 1 shows the process of how the mel-frequency cepstral coefficient features are extracted.

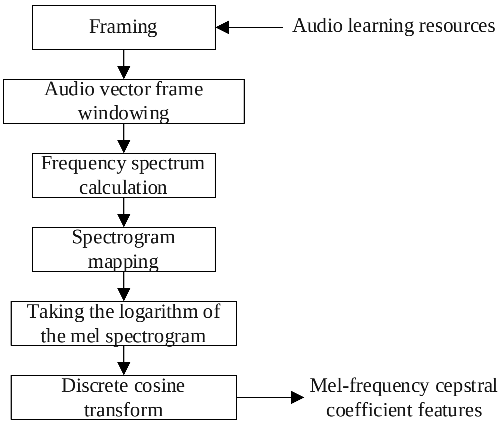

Fig. 1. Extraction process of mel-frequency cepstral coefficient features

First, convert the audio formats in the audio learning resource set to the *WAV* format and reduce the compression ratio. Use *OpenSmile* to process audio learning resources, set the feature name in the configuration file *conf*, and set the audio sampling rate to *22KHz*.

Next, set the frame length and frame displacement of the audio vector to 2048 and 512, and window the audio vector frames to eliminate the discontinuity between frames. Use the Hamming window. Eq. (1) shows the implementation process:

$$a(k)=\sum_{-\infty}^{+\infty}\psi\left[b(i)*q(k-i)\right] \tag{1}$$

where,

$$q(k)=\begin{cases}0.54-0.46cos\left(2\pi k/K-1\right),0\le k\le K-1\\0,other\end{cases} \tag{2}$$

Calculate the spectrum of each frame of the audio vector based on the short-time Fourier transform. Since the distribution of audio energy on the spectrogram is obvious, the fast Fourier transform is adopted here, which is given by Eq. (3):

$$A(j)=\sum_{i=0}^{K=1}a(k)Q^{i,j},\ j=0,1,2,K,K-1 \tag{3}$$

Map the generated spectrogram using a mel filter to obtain a mel spectrogram. The mel filter selected here is composed of *n* triangular band-pass filters. Eq. (4) shows the expression of the transfer function:

$$R[n]=\sum_{l=0}^{M-1}\left|A_x[l]\right|^2 G_n[l], 0<n<N \tag{4}$$

where,

$$G_n[l]=\begin{cases} 0, other \\ \dfrac{g(n+1)-l}{g(n+1)-g(n)}, g(n)<l<g(n+1) \\ \dfrac{l-g(n-1)}{g(n)-g(n-1)}, g(n-1)<l<g(n) \end{cases} \tag{5}$$

In order to ensure the stability of the audio signals and facilitate the calculation, take the logarithm of the spectrogram of each audio signal after obtaining the mel spectrograms of the audio learning resources. Eq. (6) shows the specific calculation process:

$$D_i=\sum_{n}^{N-1} ln\left|R[n]\right| cos\left(\frac{k\pi(i+0.5)}{N}\right) 0\le n\le N \tag{6}$$

Finally, based on the discrete cosine transform, disperse the redundant data of the audio signal spectrograms, so as to adjust the distribution of the spectrum data and concentrate the audio features in the low frequency region, and then obtain the mel-frequency cepstral coefficient features of the audio learning resources.

3 Selection of audio learning resources based on big data analysis

If the similarity between audio learning resources is calculated only based on the mel-frequency cepstral coefficient features, the accuracy of the selection results by the recommendation system cannot be improved as there are many types of features. This paper fully considered the interest degrees of the neighbouring students, and expanded the established interest degree vector of the target students with respect to music learning based on the conventional content-based audio recommendation algorithm. Figure 2 shows the flow chart of the audio learning resource selection system. The specific steps of the algorithm are described in detail as follows:

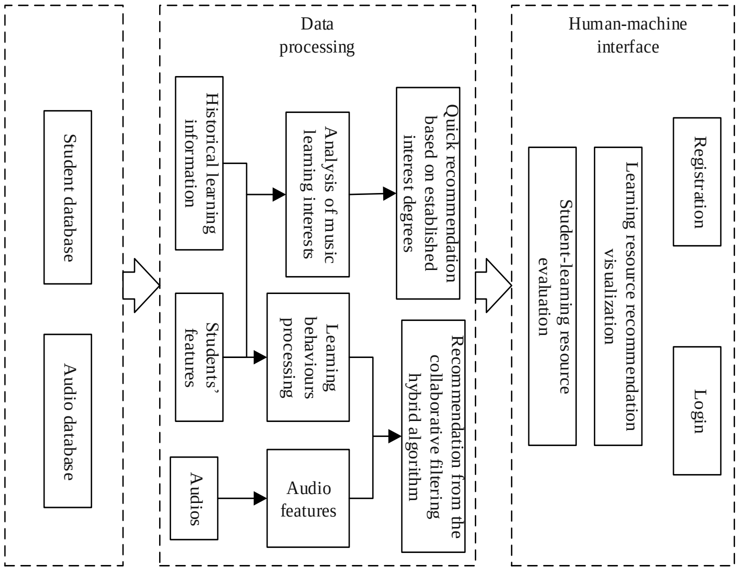

Fig. 2. Flow chart of the audio learning resource selection system

Establish a structured audio learning resource feature matrix based on the mel-frequency cepstral coefficient features of audio learning resources. Assuming that there are *n* audio learning resources and *m* feature values, Eq. (7) shows the $n \times m$-dimensional feature matrix constructed:

$$D_{nm} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & \cdots & d_{1m} \\ d_{21} & d_{22} & d_{23} & \cdots & d_{2m} \\ d_{31} & d_{23} & d_{33} & \cdots & d_{3m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ d_{n2} & d_{n2} & d_{n3} & \cdots & d_{nm} \end{bmatrix} \tag{7}$$

Based on the existing massive data of students' historical audio listening or downloading behaviours, establish the structured student interest degree matrix. Firstly, based on the student-audio frequency information, obtain the audio learning resources whose current frequency value is greater than 50%, and also the information on the mel-frequency cepstral coefficient features of these audio learning resources one by one. Through the analysis and summary of the features of audio learning resources, this paper corresponded the mel-frequency cepstral coefficient features to language, scene, genre and mood and performed cluster analysis according to the four feature attributes. Figure 3 shows the specific content of the feature attributes of audio learning resources.

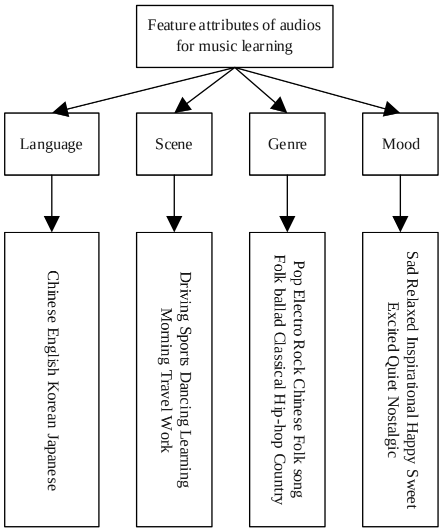

Fig. 3. Feature attributes of audio learning resources

Suppose that student v's degree of interest in feature attribute j is represented by d_{vj}, that the feature attribute appears for $|j|$ times, that the students' preference set for audio learning resources is represented by D_t, and that all feature attributes appear for $|D_j|$ times. Count the number of times the same feature information has appeared, and calculate the student's degree of interest in each feature attribute according to the interest degree calculation formula shown in Eq. (8):

$$d_{vj} = \frac{\sum_{j \in D_t} |j|}{|D_j|} \tag{8}$$

Define the student's current music learning interest vector as the set of the first 25 feature attributes of audio learning resources that have appeared for more than 2 times.

In order to effectively solve the problem that the content-based selection results of audio learning resource are not diversified, this paper incorporated the interest degrees of the neighbouring students and expanded the interest degree matrix, and based on the *top-k* similar interest values, predicted the degree of interest in the feature attributes of audio learning resources which the target student did not show interest in.

Let the student v's interest degree vector with respect to music learning be represented by $A_v=\{d_{v1},d_{v2},d_{v3}....,d_{vm}\}$, and student u's interest degree vector by $A_u=\{d_{u1},d_{u2},d_{u3}....,d_{um}\}$. Map A_v and A_u to two points in the m-dimensional space. Characterize the interest similarity between v and u with the Euclidean distance e. The smaller e is, the more similar the interest degree vectors are. Suppose that both students v and u have interest in feature attribute i of music learning resources, that student v's degree of interest in feature attribute i is represented by d_{vi}, and that student u's degree of interest in feature attribute i by d_{ui}. Eq. (9) gives the formula for calculating the m-dimensional Euclidean distance:

$$SIM(v,u) = \frac{1}{1+\sqrt{\sum_{i=1}^{m}(d_{vi}-d_{ui})^2}} \tag{9}$$

Based on the similarity between the music learning interests of students v and u and student u's interest degree vector of audio learning resource features, predict student v's degree of interest in the feature attributes of audio learning resources which he has not shown interest in. Suppose that u belongs to student v's neighbouring student set with *top-k* similar interest values, that the average interest degree of student u is denoted as λ'_u, and that the average interest degree of student v as λ'_v. The prediction formula is given as Eq. (10):

$$MUP(v,j) = \lambda'_v + \frac{\sum_{u\in TK(v)} SIM(v,u)(d_{uj}-\lambda'_u)}{\sum_{u\in TK(v)} |SIM(v,u)|} \tag{10}$$

Through the calculation by the above formula, the mixed interest degree vector of student v, which has incorporated the music learning interest degrees of the neighbouring students with similar interest values, can be obtained.

Then calculate the matching degree between the feature attribute matrix of audio learning resources and the mixed interest degree vector. The higher the matching degree, the higher the probability that students will like the audio learning resources. Suppose that the degree of matching between the interest degree of student v and the feature attributes of audio learning resource i is represented by $cos(v,i)$, that student v's degree of interest in feature attribute j by d_{vj}, and that whether the audio learning resource i contains feature attribute j by d_{ij}. The matching degree of the two is calculated as the cosine similarity shown in Eq. (11):

$$COS(v,i) = \frac{\sum d_{vj} * d_{ij}}{\sqrt{\sum d^2_{vj}} * \sqrt{\sum d^2_{ij}}} \tag{11}$$

Finally, sort the matching degrees between the feature attribute matrix of audio learning resources and the mixed interest degree vector from high to low, to obtain the selected list of audio learning resources.

In order to further improve the accuracy and stability in the prediction of students' interest in music learning, this paper proposed a collaborative filtering hybrid algorithm for audio learning resources that incorporates the interest degrees of neighbouring students. This paper expanded the student-audio learning resource rating matrix based on the conventional collaborating filtering method and then used the algorithm in the previous section to optimize the selection of learning resources. Figure 4 shows the flow chart of the collaborative filtering hybrid algorithm for audio learning resources. The algorithm consists of three levels: Level 1 is the data sets; Level 2 is to establish the interest degree and resource rating model according to the needs of the data sets and the feature extraction algorithm; and Level 3 is the execution stage of the algorithm.

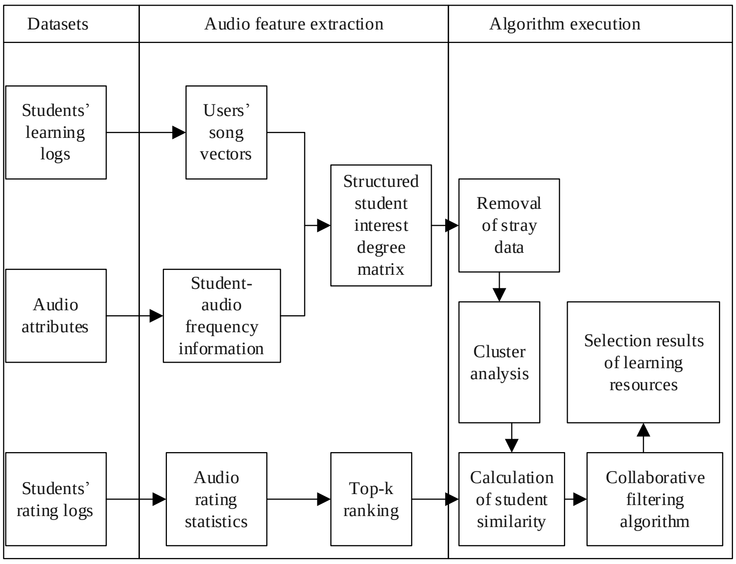

Fig. 4. Flow chart of the collaborative filtering hybrid algorithm for audio learning resources

Based on the known student-audio learning resource rating matrix, establish the following feature attribute matrix of audio learning resources:

$$D_{nm} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 1 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{12}$$

Suppose the feature attribute vector of audio learning resource i is represented by $A_i=\{d_{i1},d_{i2},d_{i3},...,d_{im}\}$, that the feature attribute vector of audio learning resource j by $A_j=\{d_{j1},d_{j2},...,d_{jm}\}$, that the set of common feature attributes of audio learning resources i and j by l, that whether audio learning resource i contains feature attribute l by $d_{i,d}$, and that whether audio learning resource j contains feature attribute l by d_{jl}, then the similarity between i and j can be calculated according to the m-dimensional Euclidean distance formula shown in Eq. (13):

$$SIM(i,j) = \frac{1}{1+\sqrt{\sum_{l=1}^{n}\left(d_{il}-d_{jl}\right)^2}} \tag{13}$$

Through the calculation by the above formula, the *top-k* audio learning resources similar to the audio learning resource *i* can be obtained. Based on the learning resource rating matrix and the similarity of learning resources, the students' ratings of audio learning resources are predicted. Suppose that the set of learning resources rated by student *v* is represented by *ER*(*v*), that the average rating obtained by audio learning resource *i* by λ'_i, that the average rating obtained by audio learning resource *j* by λ'_j and that student *v*'s rating of audio learning resource *j* by s_{vj}. The prediction method is given in Eq. (14):

$$MSP(i,v) = \lambda'_i + \frac{\sum_{j\in ER(v)} SIM(i,j) * \left(s_{vj} - \lambda'_j\right)}{\sum_{j\in ER(v)} SIM(i,j)} \tag{14}$$

After the predicted ratings of the audio learning resources by the students are obtained, transpose the matrix to obtain an expanded matrix of the original student-audio learning resource rating matrix.

Then calculate the similarity between the music learning interests of students according to the updated expanded rating matrix. Suppose that the set of audio learning resources both students *v* and *u* have rated is denoted as $QU_v \cap QU_u$, and that the average of all ratings of student *u* as λ_u. The similarity should be calculated according to the Pearson's similarity calculation method shown in Eq. (15):

$$SIM(v,u) = \frac{\sum_{l\in QU_v \cap QU_u} \left(s_{vl} - \lambda_v\right)\left(s_{ul} - \lambda_u\right)}{\sqrt{\sum_{l\in QU_v \cap QU_u} \left(s_{vl} - \lambda_v\right)^2} \sqrt{\sum_{l\in I_v \cap I_u} \left(s_{ul} - \lambda_u\right)^2}} \tag{15}$$

After the similarity between the music learning interests of *v* and *u* is calculated through the above formula, the set of neighbouring students with music learning interests similar to those of the target student can be obtained. Sort the similarities of students' music learning interests in the descending order, and then the *top-k* neighbouring students who have learning resources to recommend can be obtained.

After the *top-k* neighbour students who have learning resources to recommend are found, predict the rating of the target learning resource by the target student according to the ratings given by the neighbour students. Suppose that the set of students who have rated audio learning resource *j* among the *top-k* similar neighbouring students of student *v* is represented by $T_v(j)$, and that the rating of audio learning resource *j* by similar student *u* by s_{uj}, the prediction formula is given as follows:

$$SSP(v,j) = \lambda_v + \frac{\sum_{u\in T_v(j)} SIM(v,u)\left(s_{uj} - \lambda_u\right)}{\sum_{u\in T_v(j)} \left|SIM(v,u)\right|} \tag{16}$$

The proposed algorithm involves the similarity calculation of the large-scale feature attribute matrix, which will consume a lot of resources and time during its execution. Therefore, distributed computing based on big data analysis was adopted in this paper to improve the efficiency of resources selection. Figure 5 shows the distribution diagram of the algorithm execution.

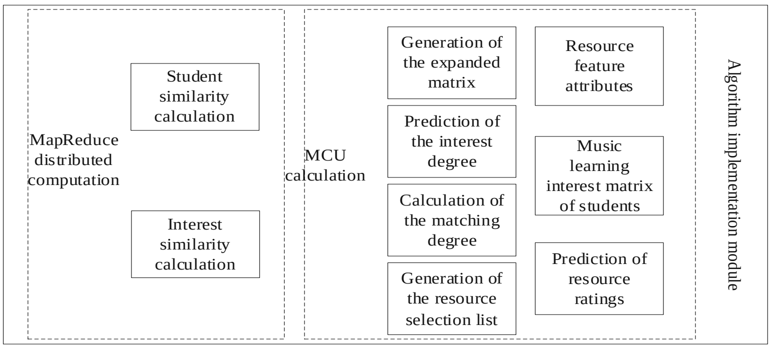

Fig. 5. Algorithm execution distribution diagram

4 Experimental results and analysis

The proposed collaborative filtering hybrid algorithm for audio learning resources incorporating neighbouring students' interest degrees was compared with the conventional content-based audio recommendation algorithm in terms of accuracy, recall rate, F1 value and satisfaction degree. The comparison results of the two algorithms in terms of accuracy, recall rate and F1 value under different numbers of selected audios are shown in Figures 6, 7, and 8. It can be seen that, compared with those of the conventional algorithm, the 3 performance indicators of the proposed algorithm were slightly lower, because the proposed method basically retained the student's original music preferences and also integrated the interest values of the neighbouring students. This algorithm solved the non-diversified selection results of the conventional algorithm at the price of a slight reduction in selection accuracy, and the experimental results were satisfactory.

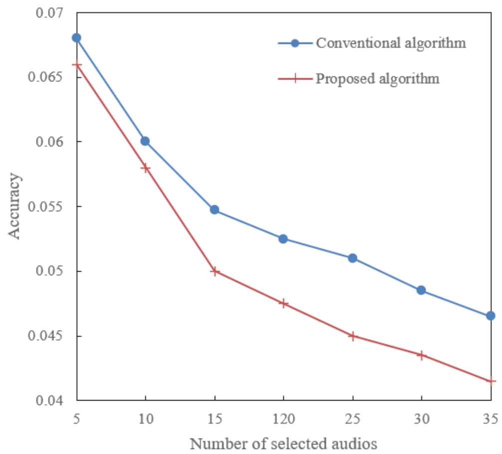

Fig. 6. Comparison of accuracy

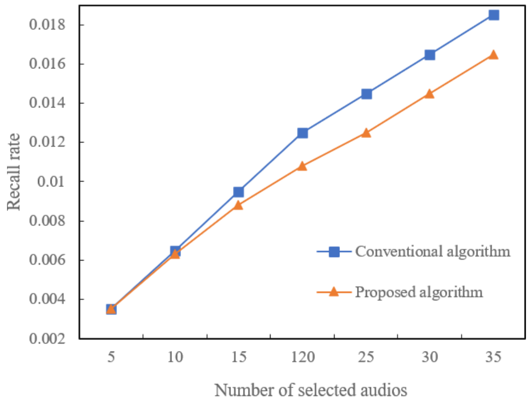

Fig. 7. Comparison of recall rates

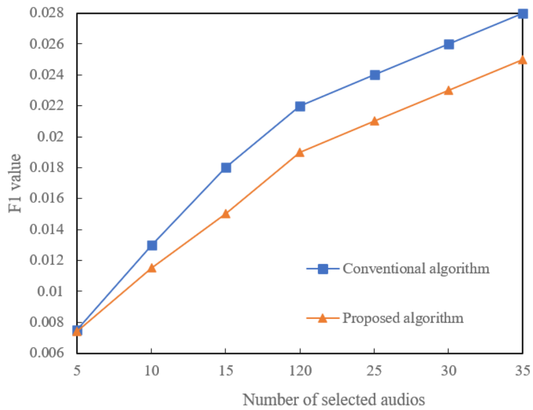

Fig. 8. Comparison of F1 values

It can be seen from the students' satisfaction degrees shown in Figure 9 that, when the number of selected audios is constant, the proposed algorithm can make students more satisfied than the conventional one, and the satisfaction degree shows a steady growth trend with the increase of the number of selected audios. The experimental results indicate that, the proposed algorithm can improve students' satisfaction while not reducing the accuracy of resource selection significantly.

It can be seen from the errors of the algorithm shown in Figure 10 that, as the sparsity of the student-audio learning resource rating matrix decreases, the root mean square error value of the prediction made by the proposed algorithm gradually decreases, which means the prediction accuracy of the algorithm increases.

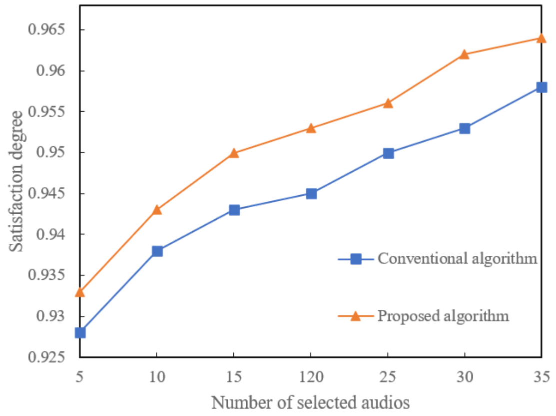

Fig. 9. Comparison of satisfaction degrees

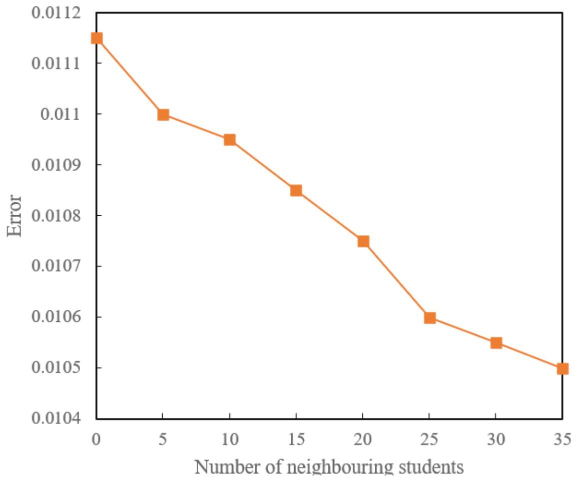

Fig. 10. Errors of the algorithm

In this paper, the conventional content-based recommendation algorithm, the content recommendation algorithm incorporating neighbours' interest degrees, the conventional collaborative filtering algorithm and also the proposed algorithm were compared in terms of F1 value, coverage, students' satisfaction and error.

From Table 1, it can be seen that the F1 value of the proposed algorithm is slightly lower than those of the others, but the proposed algorithm has obvious advantages in coverage and novelty. As the selection of audio learning resources in this paper does not simply rely on students' ratings of learning resources, but calculates the similarity of resources and students based on the feature attributes of audio signals, there is no cold boot problem as in the conventional algorithms.

Table 1. Performance comparison of different algorithms

Recommendation algorithm	Content-based recommendation algorithm	Algorithm incorporating neighbours' interest degrees	Conventional collaborative filtering algorithm	Proposed algorithm
F1 value	0.02715	0.02315	0.01958	0.01648
Coverage	0.251	0.2653	0.2057	0.5718
Satisfaction	0.917	0.938	0.902	0.982
Error	0.01748	0.01326	/	0.01748

5 Conclusions

This paper studied a method for selecting audio learning resources based on the big data of education. First, the audio signals were converted into mel spectrograms, and accordingly, the mel-frequency cepstral coefficient features of audio learning resources were obtained. Then, on the basis of the conventional content-based audio recommendation algorithm, the established interest degree vector of target students with respect to music learning was expanded. In order to effectively improve the accuracy and stability in the prediction of students' interest in music learning, a collaborative filtering hybrid algorithm for audio learning resources was proposed, which incorporates the interest degrees of neighbouring students. The experimental results show the performance of different algorithms in terms of accuracy, recall rate, F1 value, and students' satisfaction under different numbers of selected audios. According to the comparison results, the proposed algorithm has a lower F1 value than the others, but has obvious advantages in coverage and novelty. This shows that this algorithm solved the non-diversified selection results of the conventional algorithms at the price of a slight reduction in selection accuracy, and the experimental results were quite satisfactory.

6 References

[1] Peng, J. (2020). Intelligent technology-based improvement of teaching ability of professional courses in art design. International Journal of Emerging Technologies in Learning, 15(23): 193-207. https://doi.org/10.3991/ijet.v15i23.19029

[2] Tang, Y., Liang, J., Hare, R., Wang, F.Y. (2020). A personalized learning system for parallel intelligent education. IEEE Transactions on Computational Social Systems, 7(2): 352-361. https://doi.org/10.1109/TCSS.2020.2965198

[3] Baba, A.F., Cin, F.M., Ordukaya, E. (2015). Intelligent fuzzy assessment system for English Academic writing in engineering education. International Journal of Engineering Education, 31(1(A)): 83-93. https://eprints.lancs.ac.uk/id/eprint/88487

[4] Touimi, Y.B., Hadioui, A., El Faddouli, N., Bennani, S. (2020). Intelligent chatbot-LDA recommender system. International Journal of Emerging Technologies in Learning, 15(20): 4-20. https://doi.org/10.3991/ijet.v15i20.15657

[5] Baker, R.S. (2014). Educational data mining: An advance for intelligent systems in education. IEEE Intelligent systems, 29(3): 78-82. https://doi.org/10.1109/MIS.2014.42

[6] Bais, H., Machkour, M. (2019). Method and apparatus for querying relational and XML database using French language. Revue d'Intelligence Artificielle, 33(6): 393-401. https://doi.org/10.18280/ria.330601

[7] Joukhadar, A., Ghneim, N., Rebdawi, G. (2021). Impact of using bidirectional encoder representations from transformers (BERT) models for Arabic dialogue acts identification. Ingénierie des Systèmes d'Information, 26(5): 469-475. https://doi.org/10.18280/isi.260506

[8] Intayoad, W., Kamyod, C., Temdee, P. (2020). Reinforcement learning based on contextual bandits for personalized online learning recommendation systems. Wireless Personal Communications, 115(4): 2917-2932. https://doi.org/10.1007/s11277-020-07199-0

[9] Houari, H., Guerti, M. (2020). Study the influence of gender and age in recognition of emotions from algerian dialect speech. Traitement du Signal, 37(3): 413-423. https://doi.org/10.18280/ts.370308

[10] Lin, C.C., Liu, Z.C., Chang, C.L., Lin, Y.W. (2018). A genetic algorithm-based personalized remedial learning system for learning object-oriented concepts of Java. IEEE Transactions on Education, 62(4): 237-245. https://doi.org/10.1109/TE.2018.2876663

[11] Sun, J.C.Y., Yu, S.J. (2019). Personalized wearable guides or audio guides: An evaluation of personalized museum guides for improving learning achievement and cognitive load. International Journal of Human-Computer Interaction, 35(4-5): 404-414. https://doi.org/10.1080/10447318.2018.1543078

[12] Li, X., Gong, Y., Liang, Y., Wang, L.E. (2021). Personalized federated learning with semisupervised distillation. Security and Communication Networks, 2021: Article ID 3259108. https://doi.org/10.1155/2021/3259108

[13] Zhang, P., Xiong, F., Leung, H., Song, W. (2018). FunkR-pDAE: Personalized project recommendation using deep learning. IEEE Transactions on Emerging Topics in Computing, 9(2): 886-900. https://doi.org/10.1109/TETC.2018.2870734

[14] Li, H., Zhong, Z., Shi, J., Li, H., Zhang, Y. (2021). Multi-objective optimization-based recommendation for massive online learning resources. IEEE Sensors Journal, 21(22): 25274-25281. https://doi.org/10.1109/JSEN.2021.3072429

[15] Eghtesad, S. (2018). Authentic online resources for learning French. In 2018 12th Iranian and 6th International Conference on e-Learning and e-Teaching (ICeLeT), 007-012. https://doi.org/10.1109/ICELET.2018.8586757

[16] Cadzow, A. (2017). Impact of cognitive learning disorders on accessing online resources. In International Conference on Universal Access in Human-Computer Interaction, 10279: 363-381. https://doi.org/10.1007/978-3-319-58700-4_30

[17] Dai, Y., Xu, J. (2021). Study of online learning resource recommendation based on improved BP neural network. International Journal of Embedded Systems, 14(2): 101-107. https://doi.org/10.1504/IJES.2021.113834

[18] Dibie, O., Sumner, T., Maull, K.E., Quigley, D. (2016). Exploring social influence on the usage of resources in an online learning community. In EDM, 585-586.
[19] Zhang, J., Kong, D. (2021). Application and optimization of algorithm recommendation in mobile audio APP. In Journal of Physics: Conference Series, 1848(1): 012025. https://doi.org/10.1088/1742-6596/1848/1/012025
[20] Ding, H., Huang, J., Cao, H., Liu, Y. (2016). Improving cold music recommendation through hierarchical audio alignment. In 2016 IEEE International Symposium on Multimedia (ISM), 77-82. https://doi.org/10.1109/ISM.2016.0023
[21] Shepstone, S.E., Tan, Z.H., Jensen, S.H. (2014). Using audio-derived affective offset to enhance tv recommendation. IEEE Transactions on Multimedia, 16(7): 1999-2010. https://doi.org/10.1109/TMM.2014.2337845
[22] Font, F., Serra, J., Serra, X. (2014). Class-based tag recommendation and user-based evaluation in online audio clip sharing. Knowledge-Based Systems, 67: 131-142. https://doi.org/10.1016/j.knosys.2014.06.003
[23] Zhang, J., Kong, D. (2021). Application and optimization of algorithm recommendation in mobile audio app. In Journal of Physics: Conference Series, 1848(1): 012025. https://doi.org/10.1088/1742-6596/1848/1/012025
[24] Li, T. (2021). Selection of audio materials in college music education courses based on hybrid recommendation algorithm and big data. In Journal of Physics: Conference Series, 1774(1): 012019. https://doi.org/10.1088/1742-6596/1774/1/012019
[25] Bogdanov, D., Haro, M., Fuhrmann, F., Xambó, A., Gómez, E., Herrera, P. (2013). Semantic audio content-based music recommendation and visualization based on user preference examples. Information Processing & Management, 49(1): 13-33. https://doi.org/10.1016/j.ipm.2012.06.004
[26] Deng, J.J., Leung, C. (2012). Emotion-based music recommendation using audio features and user playlist. In 2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012), 796-801.
[27] Bozzon, A., Prandi, G., Valenzise, G., Tagliasacchi, M. (2008). A music recommendation system based on semantic audio segments similarity. Proceeding of Internet and Multimedia Systems and Applications-2008, 182-187.

7 Authors

Peng Wang, graduated from Odessa National Conservatory of music, Ukraine with a doctor's degree in Music art, is a lecturer working on Cangzhou Normal University. His research interests are Vocal music practice teaching and Research on practical guidance of music stage art. He has published 1 international journal, 1 national core journal and 4 school level journals of Odessa National Conservatory of music in Ukraine. During his study period, he has participated in vocal music competitions for many times and achieved excellent results. He has organized and participated in large-scale concerts for many times (email: wangpeng1232021@126.com).

Xia Wang, graduated from Hebei University of Normal with a Bachelor's degree in Musicology, is an associate professor working on Cangzhou Normal University. Her research interests are Vocal music practice teaching and Research on historical materials of Grand Canal music culture. She has participated in a national project, and presided over 2 provincial projects, published 4 core journal papers and 2 provincial journal papers. She has been rated as an excellent teacher of the year for many times (email: wx19720513@126.com).

Xia Liu, graduated from Shijiazhuang University of Railway Engineering with a master's degree in Computer application. She is an associate professor working on Cangzhou Normal University. Her research interests is computer application. She has presided over 2 provincial projects, and published 5 core journal papers and 3 provincial journal papers. She has been rated as an excellent teacher of the year for many times (email: qychmlx@163.com).

Article submitted 2022-01-07. Resubmitted 2022-02-10. Final acceptance 2022-02-11. Final version published as submitted by the authors.