International Journal of Web Applications

Print ISSN: 0974-7710
Online ISSN: 0974-7729

DLINE JOURNALS

IJWA 2024; 16 (4)
https://doi.org/10.6025/ijwa/2024/16/4/143-150

Application of Hybrid Recommendation Algorithm Based on Collaborative Filtering and Content in Cloud Vocal Music Teaching

Shenghuan Zhang[1], Lihong Xu[2*]
[1]Music College, Jimei University, Xiamen, 361021, China

[2]Minjiang Teachers College, Fuzhou,350108, China
18959262981@163.com

ABSTRACT

This article aims to introduce a new type of cloud music education platform which can meet the current high demand for music education while providing convenient, economical, and personalized services. The platform uses the Scrapy crawler framework, Web frontend, and Laravel framework and achieves efficient services through data interaction and embedded development. Using a hybrid recommendation algorithm, we can more accurately estimate the association between users and specified objects and decide whether to use this method by comparing their similarities. Through this work, we have achieved significant results in building a complete cloud music course experience, including a prediction accuracy of 8.41%, a recommendation effect of 10.21%, and a preference classification prediction accuracy of 2.55%.

Received: 9 June 2024
Revised: 1 August 2024
Accepted: 17 August 2024

Keywords: Hybrid Optimization Algorithm, Laravel Technology, Scrapy Crawler, Collaborative Filtering Technology

1. Introduction

Cloud vocal music teaching has become increasingly important in today's society and is now seen as a necessary form of entertainment. According to the "2023 China Online Music Development Thematic Research Report" by iiMedia Research, the overall sales of the music industry are expected to exceed RMB 17.58 billion in 2023, with its user base projected to exceed 570 million. With the development of the global economy, more and more consumers are beginning to accept and use a variety of innovative music courses. These courses cover traditional art forms, such as cloud vocals and jazz drums, and emerging forms of entertainment, such as esports and social media[1]. These emerging courses offer more affordable prices and enable more consumers to have a better music experience. The music teaching system based on a hybrid recommendation algorithm of collaborative filtering and content has become a hot field currently, recommending songs that meet user needs and preferences by analyzing user listening habits and song characteristics[2].

The hybrid recommendation algorithm of collaborative filtering and content can provide suitable music for users to choose from and improve the quality of online music services. It realizes personalized song recommendations by mining users' music preferences based on tags and user characteristics. However, the educational value of these music songs is not high because they do not provide precise timbre and pitch correction functions nor support professional music theory training. Considering the difficulty of obtaining user personal information, this paper proposes a music recommendation method based on music metadata and collaborative filtering based on traditional recommendation technology. This method can carry out various cloud vocal music teachings without using user personal information[3].

2. Related Work

With the development of Internet technology, all sectors have made leaps and bounds of progress relying on the Internet, and the convenience brought by the Internet is increasingly favoured by everyone. Meanwhile, online vocal music teaching has gradually grown, and more and more vocal music learners have joined the ranks of watching online teaching videos. With the transformation of education and the unfolding of society, the traditional vocal music teaching model can no longer adapt to the needs of today's society. A mature and good teaching model is becoming increasingly important, and the burgeoning online teaching model that has been thriving in recent years is widely adopted in the education sector[4]. The emergence of various platforms for online vocal music teaching excites everyone, especially MOOCs. Many teaching videos are recorded by famous vocalists, with abundant teaching resources emerging in large numbers. Some WeChat subscription accounts and short video teachings on Kuaishou and Douyin have also been developed. Learners are no longer limited to traditional face-to-face learning, and the emergence of online vocal music teaching has ushered in a new development opportunity, bringing learners many advantages while continuously innovating its own model, which is a great convenience for learners' learning[5]. It is undeniable that the emergence of this new online vocal music teaching model not only captivates the hearts of many learners but also provides good learning advantages for many universities and training institutions.

Social network information can mine potential user value more than other auxiliary information, so many algorithms based on social networks are proposed[6]. The common method is to use a trust evaluation model to calculate direct trust relationships between users to improve recommendation accuracy. Trust-based overall RSTEl2 and trust propagation-based SocialMF have achieved certain results[7]. Researchers continue to explore indirect trust mechanisms. For example, Zhang proposed a trust recommendation scheme based on evaluation values, anonymity, and the fusion of the Pareto model and confidence. However, these schemes did not fully consider asymmetric trust, and their recommendation results did not reach the ideal level when dealing with large-scale datasets.

To improve recommendation accuracy and efficiency under large data volumes, cluster-based recommendation algorithms, such as K-means clustering, fuzzy C-means method, etc. are proposed. Yun W U and others applied the fuzzy C-means clustering method to the recommendation algorithm. According to the user's category, search for neighbors with high user membership, forming a community with similar interests[8]. ARANS is a social network algorithm based on an adaptive neighbor selection mechanism proposed by Lv and others. Tian and others combined overlapping community discovery with recommendation algorithms, calculated the acquaintance between users and the relevance of users and communities, and used indirect social relationships to improve recommendation effects [9]. However, the above algorithms deal with individual users. Although the methods proposed can alleviate certain data sparsity, they have limited improvement in recommendation performance.

A collaborative filtering recommendation algorithm (CSIT-CF) that integrates community structure and user implicit trust is proposed. The algorithm first uses a trust matrix to mine the asymmetric relationship of users as trust and being trusted, calculates the user's implicit trust, and combines the implicit trust to discover the user community. Finally, it calculates the target user's predicted score for the project within the community and forms a recommendation. The algorithm can effectively suppress the data sparsity of traditional collaborative filtering algorithms while maintaining a high recommendation efficiency, thereby greatly improving the accuracy of recommendations[10].

3.Recommendation algorithm based on user Collaborative filtering

3.1. Algorithm Principle

Using the "K-Nearest Neighbors" algorithm, we can extract valuable content from the preferences of K-nearest neighbors, thereby achieving effective communication between users for mutual sharing and thus improving efficiency [11,12]. The "Neighbors" algorithm enables us to extract valuable content from the preferences of K-nearest neighbors, facilitating effective user-to-user communication to enhance efficiency. Figure 1 presents the basic concept of an algorithm used to solve the problem.

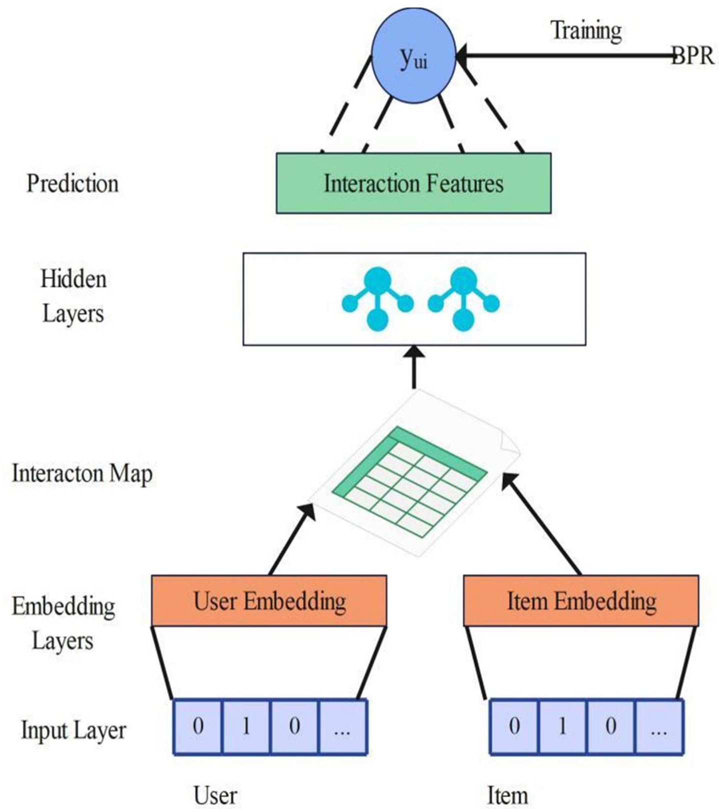

Figure 1. User-Based Collaborative Filtering Recommendation Algorithm

According to the data above, we can ascertain that the preferences of three users—A, B, and C—are roughly similar for different items. All show interest in items 1, 3, 4, etc. In addition, according to their past consumption habits, their patterns for consuming item 4 appear closely aligned. Therefore, we believe that the consumption habits of users A, B, and C towards different items are relatively similar, and user A should be prioritized for the purchase of item 4.

To more accurately identify potential customers, we propose a new method to assess users' preferences and utilize this information to evaluate the needs of other customers. The crux of this method lies in collecting extensive customer information, which is compiled into a matrix. Each matrix contains n items, each with M evaluation values. This allows us to make an accurate assessment of each customer's needs and satisfy their requirements.

The "k-nearest neighbors" search technique helps us quickly locate our desired destination. In the screening process of the "nearest neighbor" set, multiple effective metrics were adopted, including cosine similarity, Pearson correlation coefficient, k users who exceed the index limit, and so on, effectively screening the "nearest neighbor" set. We will filter out high-feedback suggestions based on audience feedback to provide superior quality service [13,14].

3.2. Introduction to Two Hybrid Recommendation Algorithms

Using the TF-IDF algorithm, we can effectively suppress the influence of the external environment, thereby obtaining accurate content recommendations. Furthermore, it can demonstrate individual user differences, providing precise analysis. Additionally, the TF-IDF algorithm can convert textual information into feature vectors, obtaining the weight of each keyword and thereby enabling accurate analysis. Through "items", we find that this collaborative filtering algorithm has high generalization ability [15,16]. It can effectively filter information related to the categories that the user initially prefers, thus reducing content confusion. "Tagged items" are effective recommendation services for specific user groups, such as music, educational resources, experts, social media, etc.

The TF-IDF algorithm can help us build a rich information model. Based on the characteristics of different items, it creates a sparse matrix with strong reliability. Also, using the collaborative filtering algorithm to compare different items can derive the optimal recommendation scheme. Finally, we can use the TOP-N analysis method to determine the best recommendation scheme based on the results of these comparisons.

The TF-IDF algorithm aims to enhance the text's readability and operability by providing a set of specific keywords. It constructs a series of weighted sparse matrices by sorting these specific keywords according to certain standards, thereby reflecting the readability and operability of specific features.

$$TF_{i,\,j} = \frac{n_{i,\,j}}{\sum_k n_{k,\,j}}$$

$$|\ DF_i = \log_{10} \frac{|\ D\ |}{\left|\left\{ j : t_i \in d_j \right\}\right| + 1}$$

In equation (2), |D| represents all texts in the corpus, $|\{j: ti \in dj\}|$ represents that each text has a specific keyword, and the value of +1 can be used to avoid the absence of a specific keyword in the text, resulting in the calculation of TF-IDF as follows:

$$TF - IDF_{ij} = TF_{i,\,j} \times IDF_i$$

According to equation (3), the higher the weight of the keywords, the more often they appear in the document, and the less frequently they occur in the entire collection of documents. In this way, we can obtain the weight sparse matrix of items.

Firstly, we should identify two items of low similarity by filtering and removing them to decrease computational complexity. We use the cosine similarity algorithm to estimate the similarity between two items, with a calculation range of [-1, 1]. Through the difference between two vectors, we can estimate their correlation, which diminishes with the distance between them. To better assess this relationship, we should calculate their correlation using previous information, such as user preferences and frequently used messages. This process allows us to filter closer messages and gather more information. Although the Pearson correlation coefficient can estimate the similarity of relationships between two users and reflect their performance at different evaluation levels, it overlooks changes in these users under different evaluation levels and variations in popular items under these levels.

When two users compare a specific set of items, their similarity is typically higher than when other users compare. However, when two users compare a specific set of items, their similarity is often lower than other pairs of users. Through Jaccard's similarity, we can measure the level of attention two individuals give to an item when using it. Then, using the quality(*u,v*) coefficient, we can decrease these attention levels to more accurately measure this level of attention. In this way, we can better understand the level of attention a user pays to an item when using it. In this context, we use the item degree di as a standard to measure a set of items. This standard is calculated based on the key attributes of the item set. We will use these key attributes to describe the performance of this set of items and use their highest values as the evaluation standard.

4. Experimental Design and Analysis

4.1. Experimental Design

In this study, we used the MovieLens dataset for comparison. This dataset contains 100 and 836 comments from 610 users on 9 and 742 movies. To determine the accuracy of these data, we divided them into 10 parts and used some as test samples, while the remaining 9 parts were used as model samples. We will repeat this process until the average score of all samples reaches a unified standard.

Through practical applications, we have found that three different algorithms can more accurately identify user preferences. These three algorithms include multi-level collaborative filtering, user opinion transmission algorithms, and the Pearson algorithm, all of which can more accurately identify user preferences. We measure the average error and Root-mean-square deviation of the algorithms and use the half-life method to measure their sorting effect. Finally, we use the Accuracy and F1 methods to test their fineness.

Through an effective prediction, TP refers to the information about user preferences provided by the algorithm, FP refers to the information about user preferences provided by the algorithm, FN refers to the information about user preferences provided by the algorithm, and TN refers to the information about user preferences provided by the algorithm. With the increase of several key factors, the algorithm can more accurately identify users' likes or dislikes, thus making decisions more effectively.

4.2. Experimental Results and Analysis

According to Figure 2, when the number of neighbors in 1 is high, the error of the CRQQ algorithm is significantly smaller than Pearson Entropy, reaching 6.68%. However, when there are many neighbors in UOS, Pearson Entropy, and MLCF, the error of the CRQQ algorithm is still small, reaching an error of 0.57% to 6.68%. After experimental verification, our proposed algorithm can quickly select the best neighbor with extremely high efficiency and achieve high prediction accuracy.

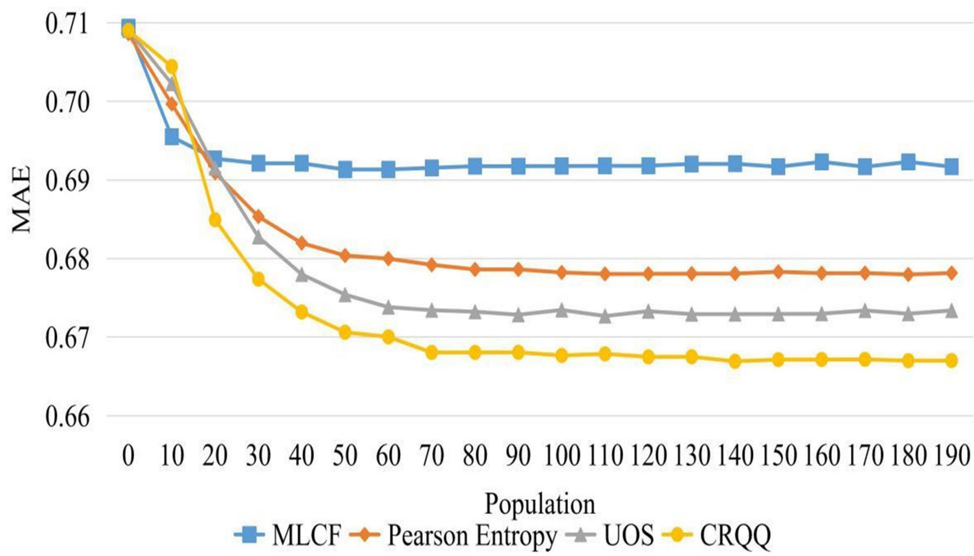

Figure 2. Comparison of MAE results for several algorithms

The application of RMSE can greatly reduce the error of the CRQQ algorithm, putting it in a leading position among various algorithms. Therefore, its prediction accuracy is extremely high, and it hardly produces any significant deviation. In addition, as the number of neighbors of the CRQQ algorithm increases, its error will dramatically decrease, with the reduction ranging from Pearson Entropy, UOS, and 1.40% to 3.24%. This leads it to a leading position among various algorithms, hence significantly enhancing its performance. After applying the CRQQ algorithm, the error of MLCF significantly decreases from 0.92% to 3.23%, indicating that the precision of the CRQQ algorithm is superior to traditional algorithms, thus effectively improving prediction accuracy.

To validate the performance of this algorithm under the same conditions, we compare the algorithm in this paper with the following three algorithms:

1) Social MF. This algorithm utilizes direct social relationships and uses friend information and trust propagation mechanisms to optimize the user feature matrix.

2) Using the Trust MF algorithm, we can simultaneously analyze the rating matrix and social matrix and transform the trust relationship between users into two independent feature vectors, thereby improving the recommendation's accuracy.

3) UCEC&CF. This algorithm corrects and dynamically evolves the scoring matrix, uses user interest to measure user relevance, and finally makes recommendations within user groups.

Figure 3 compares the recommendation effects of our algorithm with Social MF, Trust MF, and UCEC&CF under the same dataset and experimental environment conditions.

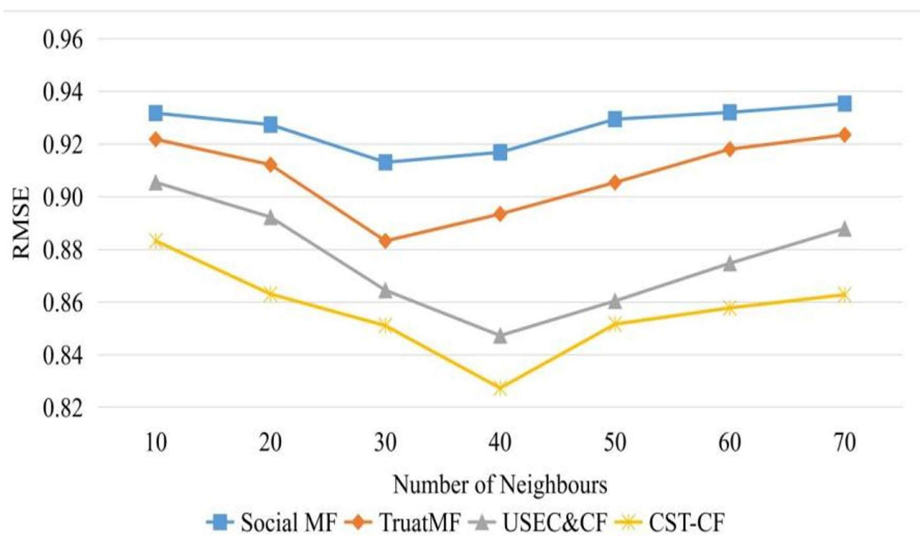

Figure 3.Comparison of RMSE metrics among various algorithms

From Figure 3, we can see that the HLU ranking ability of the four algorithms shows a good trend, performing at a high level in the early stages and experiencing some fluctuations in the later stages. However, CRQQ still maintains the first place. Specifically, when the number of neighbors reaches 26, its HLU ranking ability reaches 1.02. In a similar environment, the result of the UOS dataset is 0.99, while the result of the Pearson Entropy algorithm is 0.96, with the changes between the two not being significant. However, the later results of the MLCF algorithm are superior to the Pearson Entropy algorithm. The ranking performance of the CRQQ algorithm significantly surpasses other algorithms, ranging from 0.63% to 10.21%, indicating that it has stronger comprehensiveness, can effectively improve the algorithm's performance, and thereby more effectively achieve the expected results.

5. Conclusions

The focus of this paper is to explore how to provide high-quality audio information through a cloud audio education platform. Since most audio information is in text form, and given that audio information contains a large amount of information with a high update frequency, we must take some measures to improve the accuracy of this information. The new algorithm we propose is based on item-based collaborative filtering and content hybrid recommendation, which can help us better manage audio information and improve audio quality. Combining multiple technologies, we have successfully built a comprehensive cloud music course system that meets consumer requirements for diversified, efficient, and low-cost music courses.

References

[1] Wu, L. (2021). Collaborative Filtering Recommendation Algorithm for MOOC Resources Based on Deep Learning. *Complexity, 2021*(46), 1–11.

[2] Xu, J. (2021). An Improved Hybrid Recommendation Algorithm Based on Feature Preference. *Journal of Physics: Conference Series, 1994*(1), 01–20.

[3] Wang, X., Dai, Z., Li, H., et al. (2021). Research on Hybrid Collaborative Filtering Recommendation Algorithm Based on the Time Effect and Sentiment Analysis. *Complexity, 2021*(2), 1–11.

[4] Cai, B., Zhu, X., & Qin, Y. (2021). Parameters Optimization of Hybrid Strategy Recommendation Based on Particle Swarm Algorithm. *Expert Systems with Applications, 168*, 114388.

[5] Arik, A., Okyay, S., & Adar, N. (2021). Hybrid Course Recommendation System Design for a Real-Time Student Automation Application. *European Journal of Science and Technology, Special Issue*(26), 85–90.

[6] Zhang, Y., Liu, Z., & Sang, C. (2021). Unifying Paragraph Embeddings and Neural Collaborative Filtering for Hybrid Recommendation. *Applied Soft Computing, 106*, 107345.

[7] Yang, Y., & Ko, Y. C. (2022). Design and Application of Handicraft Recommendation System Based on Improved Hybrid Algorithm. *International Journal of Pattern Recognition and Artificial Intelligence, 2022*(2), 36.

[8] Wu, Y., Li, J., & Ma, Y. (2022). A Hybrid Music Recommendation Model Based on Personalized Measurement and Game Theory. *Chinese Journal of Electronics, 32*, 1–10.

[9] Lv, Y., & Kong, J. (2021). Application of Collaborative Filtering Recommendation Algorithm in Pharmacy System. *Journal of Physics: Conference Series, 1865*(4), 042113.

[10] Tian, P. (2021). Collaborative Filtering Recommendation Algorithm in Cloud Computing Environment. *Computer Science and Information Systems, 18*, 8.

[11] Zhang, J., Peng, Q., Sun, S., et al. (2014). Collaborative Filtering Recommendation Algorithm Based on User Preference Derived from Item Domain Features. *Physica A: Statistical Mechanics and Its Applications, 396*, 66–76.

[12] Cui, Z., Xu, X., Fei, X., et al. (2020). Personalized Recommendation System Based on Collaborative Filtering for IoT Scenarios. *IEEE Transactions on Services Computing, 13*(4), 685–695.

[13] Lytvyn, V., Vysotska, V., Shatskykh, V., et al. (2019). Design of a Recommendation System Based on Collaborative Filtering and Machine Learning Considering Personal Needs of the User. *Eastern-European Journal of Advanced Technologies, 4*(2), 6–28.

[14] Chen, J., Zhao, C., Uliji, et al. (2020). Collaborative Filtering Recommendation Algorithm Based on User Correlation and Evolutionary Clustering. *Complex & Intelligent Systems, 6*,

147–156.

[15] Subramaniyaswamy, V., Logesh, R., Chandrashekhar, M., et al. (2017). A Personalized Movie Recommendation System Based on Collaborative Filtering. *International Journal of High Performance Computing and Networking, 10*(1–2), 54–63.

[16] Wei, S., Ye, N., Zhang, S., et al. (2012). Collaborative Filtering Recommendation Algorithm Based on Item Clustering and Global Similarity. *2012 Fifth International Conference on Business Intelligence and Financial Engineering, IEEE*, 69–72.