Personalized Music Recommendation model based on Machine Learning

Dr Mohammad Gouse Galety
Department of Information Technology
Catholic University in Erbil,
Erbil, Kurdistan Region, Iraq
gousegal@gmail.com

Dr.R.Thiagarajan
Department of IT
Prathyusha Engineering College
Chennai, India
rthiyagarajantpt@gmail.com

R.Sangeetha
Department of ECE
SSM Institute of Engineering and Technology
Dindigul, India
sangeethassmece@gmail.com

L.K.Balaji Vignesh
Department of ECE
Ramco Institute of Technology
Rajapalayam, India
balagkannan@gmail.com

Dr.S.Arun
Department of ECE
Prathyusha Engineering College
Chennai, India
yesarun1810@gmail.com

Dr.R.Krishnamoorthy
Department of ECE
Prathyusha Engineering College
Chennai, India
krishnamoorthy.mkr@gmail.com

2022 8th International Conference on Smart Structures and Systems (ICSSS) | 978-1-6654-9761-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSSS54381.2022.9782288

***Abstract*—A music recommendation system suggests songs to an individual user on his preferences. There are many different sorts of music to choose from. The world of music is so vast that it is impossible to listen to all the songs one desires. As a result, we create a model that supports a user in discovering music that he could enjoy. It collects individuals who share the user's passions and picks knowledge and resemblance associations depending on the user's past. The information gathered from user evaluations is used to make suggestions. The main focus of the study is on the context-aware recommendation process's insufficient integration of context data with the emergence of new attractions. Using libraries like NumPy and Pandas, we used a library of songs to uncover connections across individuals and music so that a hit album might be offered to individuals derived from history. In addition to Count Vectorizer (CV), we'll use Cosine similarity (CS). In addition, when a piece of given music is processed, a front end with a flask will provide us with the suggested tracks.**

Keywords— Music recommendation system, NumPy, Pandas and Count Vectorizer.

I. INTRODUCTION

The Web's capabilities are expanding due to the fast expansion of the Internet and technology. On the one hand, the resources available through the Internet are growing increasingly abundant, bringing considerable comfort to people's lives. However, on the other side, the enormous data space provides users with more options. Simultaneously, users get lost in a sea of data to find the information they need, a phenomenon known as "information overload." To address these issues, a proposed recommendation system was developed, which dynamically offers objects that match users' interests based on their preferences [1-3].

Choosing a suitable recommendation system is critical to the effective implementation of a personalized recommendation engine, and the algorithm's performance directly affects recommendation quality. The most prevalent recommendation approaches are frequent pattern evaluation, sentiment classification recommendation, hybrid recommendation [4-8], and content-based recommendation technology [9-12].

In the field of recommendation, collaboration filtering suggestion is now the most widely used effective at the same time and the most active area of study in the field of recommendation algorithms. Collaborative filtering identifies how comparable individuals' interests are now, and filters and screens target consumers depending on similar ones.

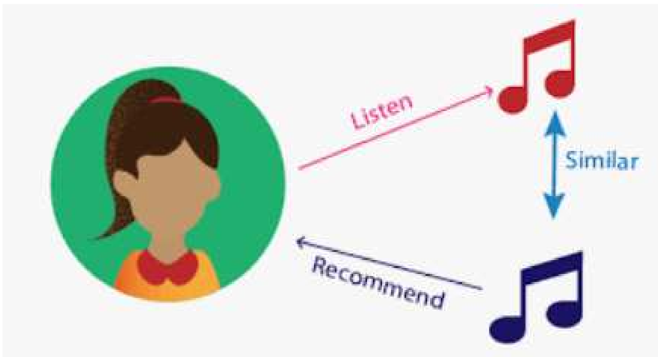


Fig. 1. Basic recommendation scheme

The primary premise of collaborative filtering would be that users share comparable values, beliefs, levels of knowledge, and curiosity inclinations, as well as a similar need for information [13 - 16]. As a result, collaborative filtering technology has a benefit over traditional methods, some topics that are not easy to analyze material for, such as abstract resource objects and personal taste. Collaborative filtering expertise can efficiently utilize the data to analyze many users with comparable interests, allowing for fewer user feedback and faster-personalized learning of hidden interests. The coordinated filtering system has piqued the interest of a growing number of academics due to its wide range of applications [17]. Meanwhile, the combined filtering technique has been elevated with research domains in ML. Collaborative filtering recommendation (CFR) and Content-based recommendation (CBR) are two researched recommendation systems. Continuously improving information filtering technology is a content-based recommendation. The system does not require users to provide feedback on projects; instead, it simply analyzes the content knowledge from existing project selections to recommend new initiatives. For most CBR engines, the content info of such a task is defined based on phrases [18 - 22].

II. LITERATURE SURVEY

Data science and ML are involved in suggestion, but music suggestion also necessitates audio signal processing and music theory. Music recommendations can draw on various disciplines, including physiology, psychology, anthropology, and philosophy. Understanding the human relationship to music is a multidisciplinary study project that spans several fields. More investigation in these areas will be required to create a system that understands music. Many researchers proposed using machine learning to analyze and forecast songs. The primary feature of this method is predicting whether users will like or dislike songs based on their collection.

The music recommendation is based on the user's preferences from their playlist as shown in Fig. 1. Algorithms like decision trees, random forests, and regression were used. The song prediction is presented as a graphical user interface (GUI), where the user can enter the information of the attribute and forecast whether the song will be liked or disliked. Currently, recommendation systems use a hybrid approach to create recommendations, exploiting substantial data pools. Methods that consider social, contextual, emotional, or cultural components of music are come out in the literature, but at rest being investigated.

J. Shi [8] proposed a solution extended to different platforms and domains. This research focuses on enhancing music recommendation algorithms. Once more variables are included, current systems are inefficient. Tunes Recommendation System is a categorization machine that uses deep learning that combines a hybrid of CFR and CBR as input to build a recommendation scheme by instantaneous prediction. We use the Spotify Recsys Challenge data to demonstrate our method, achieving precision scores of up to 88 percent at an impartial favoritism threshold.

In a personalized music recommender system, G. Yang [9] describes collaborative filtering. The user is recommended music using a content filtering suggestion system that blends output values with log data. The suggested system contains log files that preserve the user's past music playlist history. The proposed music recommendation system uses the user's log file history to provide music suggestions for every suggestion. Based on the audio, content-based techniques provide recommendations.

Mohammed et al. [10] focuses on the numerous characteristics of songs. The layout lets the customer listen to music while providing recommendations depending on the currently playing song. Firebase firestore and firebase storage are used to store the recommendations. The algorithm gives each song's acoustic features and lyrics rank. The algorithms employed are AI and KNN regression. The program calculates the score of similarity between two pieces. The user will be recommended songs with similar mean counts to the currently playing song. The primary goal of music recommendation in this study is to provide solid human-computer interaction and deliver good recommendations to users.

By merging the filtering process with Deep Learning, Y. Zhang [11] hopes to improve the RS. It will recommend new music using the traditional filtering technique and the song's album art. The hybrid Recommendation Systems will look for distinctive labels in the song's album art.

Hui Ning et al. [12] proposed a content matching approach to handle several songs and users as a suggestion scheme for search engines is becoming increasingly significant in electronic commerce and online video applications, particularly in search engines. As a result, their technology may well be used to improve user experience on search engines like Google, Bing, and Baidu.

Schedl et al. [13] to reduce the score, the forget function and memory are used, then the consumer and attributes are taken into account. To calculate similarity; second, we train feature weights using logistic regression; at last, suitable weights are being used to incorporate consumer and item pattern matching recommendation consequences. The prediction accuracy of the revised CFR algorithm in this work has substantially increased due to the preceding enhancements. The efficiency of the MRAPP algorithm is proved by preliminary theoretical proof and simulated studies.

Afchar et al. [14] is working on a prototype for dynamic music recommender systems related to human emotions. The tunes for each emotion are trained using a combination of features and ML algorithms. To keep the consumers interested, suitable music is played when the mood is decided from the source photographs. The application is linked to human emotions with this strategy, offering users a more personalized experience. As a result, our proposed system would use artificial intelligence and machine learning techniques to identify human emotions to construct an emotion-based music player. We utilise openCV to recognize emotions and make music recommendations in our experiments.

According to Adiyansjah et al. [15], the goal is to extract features from every song and compile music features. This dataset aids in the development of a model. This methodology is based on the grouping principle, which creates songs with similar characteristics. When a user listens to a song S, he gets suggested another music that has comparable attributes to S. The user's song listening history is also kept organized by genre, and a library of songs that pique his attention is constructed.

Hui et al. [16] employed the random forest method and the decision tree algorithm. Both systems seek to predict a decision based on a variety of factors. We also employed a cross-validation strategy to improve the model's precision and effectiveness. We recursively modified the training and testing data sets to achieve the best results. We had to evaluate a user's music history frequency list favorite tracks for a particular user. Then, we had to predict based on all of this data about what music the user might enjoy. We offer a musical prediction model based on content and a simple technique of ranking movies based on their approval rating in this study. Python and machine learning algorithms were used to create the system.

Urdaneta et al. [17] is developing a music recommendation system regarding feature resemblance in acoustic input. A convolutional recurrent neural network (CRNN) is used in this study to extract features, and a similarity measure is used to assess the significant positive association. As per the findings of this research, people favored recommendations on musical styles over suggestions based just on similarities.

III. SYSTEM MODEL

Amongst the several categorization approaches to be employed, the goal of this study is to figure out which DM approach is the most effective and precise. Furthermore, determining the impact of the intercity express data ratio on predictive performance.

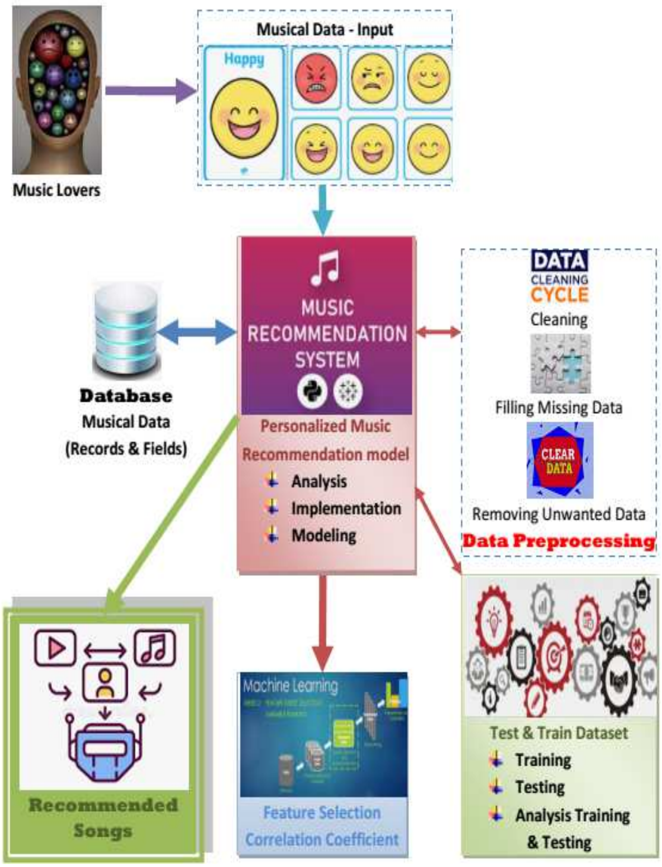


Fig. 2. System Model

Fig. 2 depicts the system's architecture. This section goes through each block in detail. Calculating the similarity measure of features extraction (equation 1) from one song to another is how the recommender system works. Because the mining characteristics are vectors, their distance can be calculated. To begin, For each genre, we picked one genre of music to act as the cornerstone for the recommendation systems. Following that, the basis music variety prediction is considered using CNN [23]. The vectors created earlier than the categorization layer are worn to build the suggestions. After the base music features have been collected, perform cosine similarity computations on different music qualities [24 -28]. Determine the resemblance between two music having a similar set of characteristics, where the first music has a vector $a = [a_1, a_2, ..., a_n]$. The second piece of music has a vector of $b = [b_1, b_2, ..., b_n]$. Here is the formula for calculating cosine similarity between two pieces of music amongst the several categorization approaches to be employed, the goal of this study is to figure out which DM approach is the most effective and precise. Furthermore, determining the impact of the intercity express data ratio on predictive performance:

$$cos\theta = \frac{\sum_{c=1}^{n} a_c b_c}{\sqrt{\sum_{c=1}^{n} a_c^2}\sqrt{\sum_{c=1}^{n} b_c^2}} \quad (1)$$

A. Data Collection and Analysis:

For the research, the real dataset is used. We used 1500 songs and 15 categories of music data, which included numerical and categorical aspects. Every column in the musical previously been linked set represents a distinct employee feature, and each existing system a single piece of musical data.

B. Pre-processing and Preparation of Data:

The process made preparing the data is performed after the process of prioritized collection. It is critical to modify the information to fit the models and deliver an enhanced domino effect. We accomplished duties such as filling in missing data cleaning, and removing unneeded data during this stage [29]. Spotify's data contained a lot of properties that were no longer relevant, i.e., didn't provide any helpful information, such as Artist, Top Genre, Title, BGM, Liveness, Energy, and so on; in this phase, these attributes were eliminated.

C. Selection of Features:

Feature selection is a fundamental idea in DM and Deep Learning. The predictor variable has a lot of columns, which is a practice of picking the required essential factors in data meant to enhance ML outcomes and make it much more reliable. As a result, the correlation coefficient is determined to see which ones are the most relevant and then employed in instructional strategies. Then we can figure out what the essential aspects are that influence effectiveness.

D. Test and Train Dataset:

Splitting input among sample and training datasets is crucial in evaluating DM models since it mitigates the consequences of inconsistent data and provides a greater understanding of the model's properties. The training set comprises all irrelevant data, while the test data has all the necessary groups for prediction. We divide the test data into erratic proportions to learn the assessment of prediction. Our study attempts to discover the most important characteristics that may have a positive effect on the accuracy of music performance estimation techniques using a variety of feature selection strategies.

IV. Experimental Results

In the scheme, the user first approaches the title track they desire; once the user has keyed the requisite song, ten equivalent music are suggested. Initially, the approach considers three primary features: Title, Artist, and Top Genre, which are calculated using Angular and Euclidean distances. Using the count vectorizer to count the number of phrases and using cosine similarity in structured data to calculate the similarity score [34]. Because we are utilizing several parameters, a function is constructed to combine the fields' data to generate the similarity score of the supplied characteristics well before the count vectorizer class records information. Any NaN cells are represented, including an empty set if they are discovered.

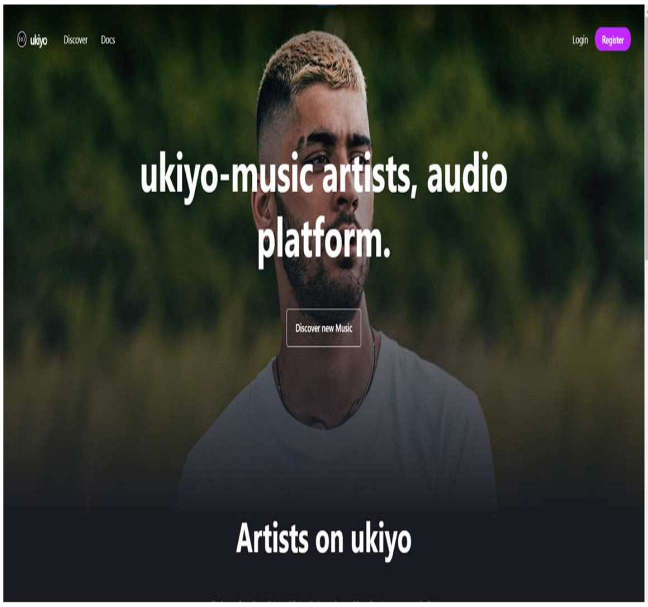


Fig. 3. Home Page

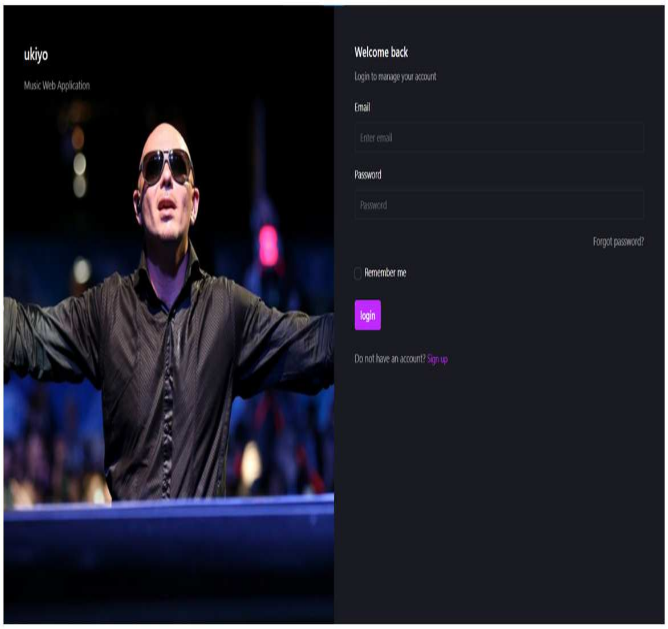


Fig. 4. Login Page

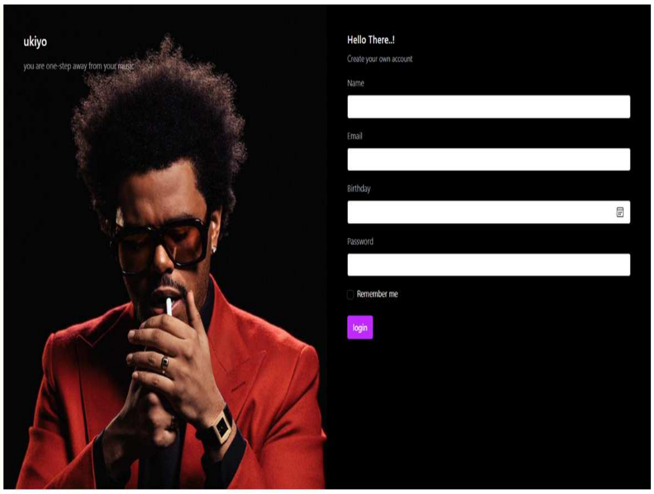


Fig. 5. Sign-in Page

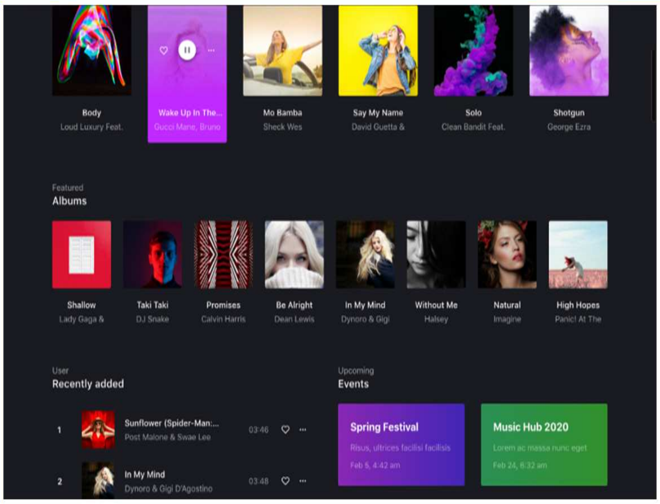


Fig. 6. Discover the Music Page

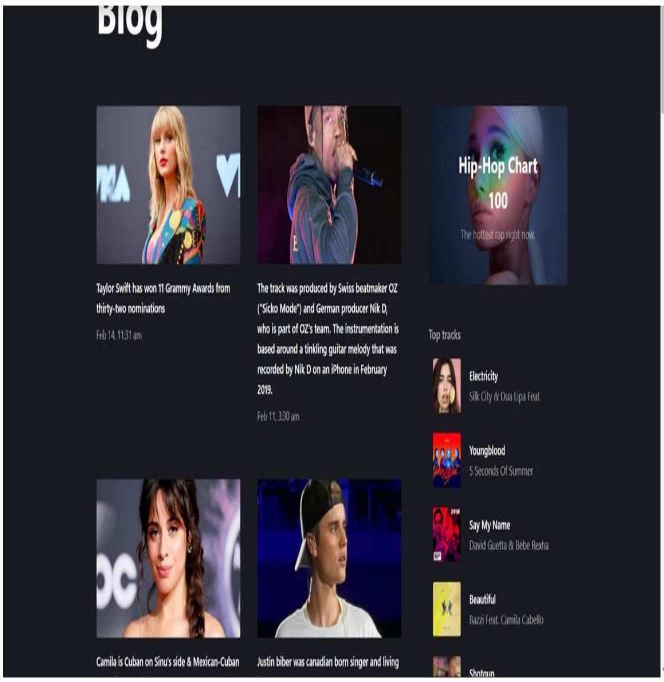
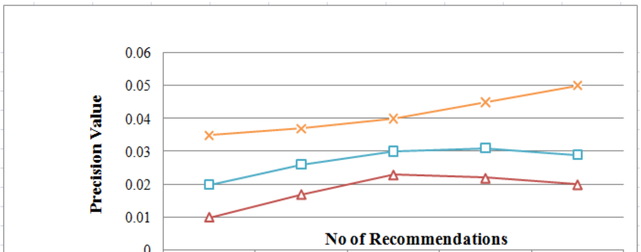
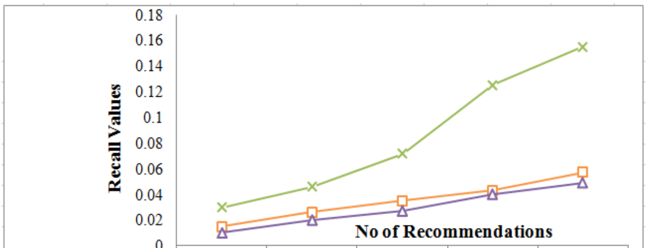


Fig. 7. Blog Page



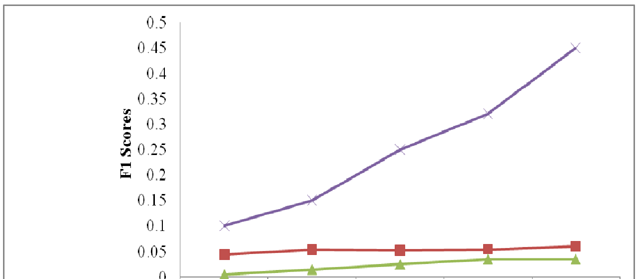
	10	20	30	40	50
CV	0.02	0.026	0.03	0.031	0.029
CS	0.01	0.017	0.023	0.022	0.02
CVCS (Proposed)	0.035	0.037	0.04	0.045	0.05

Fig. 8. Comparision of Precision Values



	10	20	30	40	50
CV	0.015	0.026	0.035	0.043	0.057
CS	0.01	0.02	0.027	0.04	0.049
CVCS (Proposed)	0.03	0.046	0.072	0.125	0.155

Fig. 8. Comparision of Recall Values



	10	20	30	40	50
CV	0.044	0.053	0.052	0.054	0.06
CS	0.005	0.015	0.025	0.035	0.035
CVCS (Proposed)	0.1	0.15	0.25	0.32	0.45

Fig. 8. Comparision of F1 Scores

V. Conclusion

Algorithms for making recommendations are an underestimated component of our daily lives, deciding which music we hear on Itunes and online. The complexity and efficiency of these techniques are still being refined through research. Our study discusses one such recommendation system that uses libraries and Cosine similarity coupled with Count Vectorizer to uncover associations among individuals and music. Depending on their previous histories, a new track might be suggested. Based on our findings, we believe that additional music elements should be included in future studies to increase the recommender system's effectiveness.

References

[1] Mohammed Hassan and Mohamed Hamada, “Improving prediction accuracy of multicriteria recommender systems using adaptive genetic algorithms,” Intelligent Systems Conference, 2017, pp. 326–330.

[2] Ruchika, Ajay Vikram Singh, and Mayank Sharma, “Building an effective recommender system using machine learning based framework,” International Conference on Infocom Technologies and Unmanned Systems, 2017, pp. 215-219.

[3] Carter Chiu and Justin Zhan, (2018) “Deep learning for link prediction in dynamic networks using weak estimators,” IEEE Access, vol. 6, pp. 35937 – 35945.

[4] J. A. Shanny and K. Sudharson, (2014) "User preferred data enquiry system using mobile communications," International Conference on Information Communication and Embedded Systems, pp. 1-5, doi: 10.1109/ICICES.2014.7033943.

[5] Yash Ketan Bhanushali and Yash Shankarbhai Patel, "Movie Recommendation System," International Research Journal of Engineering and Technology, vol. 8, no. 4, pp. 3220- 3223, 2021.

[6] Schedl Markus, "Deep Learning in Music Recommendation Systems," Frontiers in Applied Mathematics and Statistics, volume 5, 2019, doi:10.3389/fams.2019.00044.

[7] I. H. Mwinyi, H. S. Narman, K. C. Fang, and W. S. Yoo, “Predictive self-learning content recommendation system for multimedia contents,” in 2018 Wireless Telecommunications Symposium (WTS), April 2018, pp. 1–6.

[8] Juanjuan Shi, "Music Recommendation Algorithm Based on Multidimensional Time-Series Model Analysis", Complexity, vol. 2021, Article ID 5579086, 11 pages, 2021. https://doi.org/10.1155/2021/5579086

[9] Gao Yang, "Research on Music Content Recognition and Recommendation Technology Based on Deep Learning", Security and Communication Networks, vol. 2022, Article ID 7696840, 8 pages, 2022. https://doi.org/10.1155/2022/7696840.

[10] Mohammed Fadhel Aljunid, Manjaiah Dh, "An Efficient Deep Learning Approach for Collaborative Filtering Recommender System," Procedia Computer Science, Volume 171, 2020, Pages 829-836, https://doi.org/10.1016/j.procs.2020.04.090.

[11] Yan Zhang, "Intelligent Recommendation Model of Contemporary Pop Music Based on Knowledge Map", Computational Intelligence and Neuroscience, vol. 2022, Article ID 1756585, 8 pages, 2022. https://doi.org/10.1155/2022/1756585.

[12] Hui Ning, Qian Li, "Personalized Music Recommendation Simulation Based on Improved Collaborative Filtering Algorithm", Complexity, vol. 2020, Article ID 6643888, 11 pages, 2020. https://doi.org/10.1155/2020/6643888.

[13] Schedl, M., Zamani, H., Chen, CW. et al. “Current challenges and visions in music recommender systems research,” Int J Multimed Info Retr, vol. 7, 95–116 (2018). https://doi.org/10.1007/s13735-018-0154-2

[14] Afchar, Darius, Alessandro B. Melchiorre, Markus Schedl, Romain Hennequin, Elena V. Epure, and Manuel Moussallam. "Explainability in Music Recommender Systems." *arXiv preprint arXiv:2201.10528* (2022).

[15] Adiyansjah, Alexander A S Gunawan, Derwin Suhartono, "Music Recommender System Based on Genre using Convolutional Recurrent Neural Networks," Procedia Computer Science, Volume 157, 2019, Pages 99-109, doi: https://doi.org/10.1016/j.procs.2019.08.146.

[16] Hui Ning, Qian Li, and Wei Wang. 2020. Personalized Music Recommendation Simulation Based on Improved Collaborative Filtering Algorithm. Complex. 2020 (2020). DOI:https://doi.org/10.1155/2020/6643888

[17] Urdaneta-Ponte, M.C.; Mendez-Zorrilla, A.; Oleagordia-Ruiz, I. Recommendation Systems for Education: Systematic Review. Electronics 2021, 10, 1611. https:// doi.org/10.3390/electronics10141611.

[18] Y. Chen, Q. X. Zhang, and F. Akhtar, “A time effect based collaborative filtering approach for user preference statistics and recommendation,”Journal of Physics: Conference Series, vol. 1453, no. 1, Article ID 012140, 2020.

[19] Dhinakaran D., Joe Prathap P.M. (2022) Ensuring Privacy of Data and Mined Results of Data Possessor in Collaborative ARM. In: Ranganathan G., Bestak R., Palanisamy R., Rocha Á. (eds) Pervasive Computing and Social Networking. Lecture Notes in Networks and Systems, vol 317. Springer, Singapore. https://doi.org/10.1007/978-981-16-5640-8_34.

[20] Xiaoyu Tang, Yue Xu and Shlomo Geva, "Factorization-based primary dimension modelling for multidimensional data in recommender systems,"International Journal of Machine Learning and Cybernetics, vol. 10, no. 8, pp. 2209–2228, 2019.

[21] Dazhi XU. 2020. Research on music culture personalized recommendation based on factor decomposition machine. Personal Ubiquitous Comput. 24, 2 (Apr 2020), 247–257. DOI:https://doi.org/10.1007/s00779-019-01343-9

[22] J. Aruna Jasmine, V. Nisha Jenipher, J. S. Richard Jimreeves, K. Ravindran and D. Dhinakaran, "A traceability set up using Digitalization of Data and Accessibility," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), 2020, pp. 907-910, doi: 10.1109/ICISS49785.2020.9315938.

[23] Damak Khalil, Nasraoui Olfa, Sanders William Scott, "Sequence-Based Explainable Hybrid Song Recommendation," Frontiers in Big Data, Volume 4, 2021, DOI=10.3389/fdata.2021.693494.

[24] XINXI WANG, YI WANG, DAVID HSU and YE WANG, "Exploration in Interactive Personalized Music Recommendation: A Reinforcement Learning Approach", ACM Trans. Multimedia Comput. Commun. Appl., Vol. 11, No. 1, Article 7, 2014, DOI: http://dx.doi.org/10.1145/2623372.

[25] K.Sudharson, A. M. Ali and N.Partheeban, "NUITECH – Natural user interface technique foremulating computer hardware", International Journal of Pharmacy & Technology, Vol. 8, No. 4, pp. 23598-23606, 2016.

[26] M. Wang, Y. Xiao, W. Zheng, X. Jiao and C. -H. Hsu, "Tag-Based Personalized Music Recommendation," *2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN)*, 2018, pp. 201-208, doi: 10.1109/I-SPAN.2018.00040.

[27] Wei Lu, "Design of a Music Recommendation Model on the Basis of Multilayer Attention Representation", Scientific Programming, vol. 2022, Article ID 7763726, 8 pages, 2022. https://doi.org/10.1155/2022/7763726.

[28] Assuncao, W.G., Piccolo, L.S.G. & Zaina, L.A.M. Considering emotions and contextual factors in music recommendation: a systematic literature review. Multimed Tools Appl 81, 8367–8407 (2022). https://doi.org/10.1007/s11042-022-12110-z.

[29] Jingzhou Yang, "Personalized Song Recommendation System Based on Vocal Characteristics", Mathematical Problems in Engineering, vol. 2022, Article ID 3605728, 10 pages, 2022. https://doi.org/10.1155/2022/3605728.