INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 05, May 2024, pp: 387-391

e-ISSN : 2583-1062

Impact Factor: 5.725

VOICE - BASED MUSIC RECOMMENDATION SYSTEM USING MACHINE LEARNING

Sanjay A[1], Gowtham M[2], Karthikeyan M[3], Jagadeesh N[4]

[1,2,3]Student, Computer Science Engineering, Sri Venkateswaraa college of Technology, Chennai, Tamilnadu, India.

[4]Assistant Professor, Computer Science Engineering, Sri Venkateswaraa college of Technology, Chennai, Tamilnadu, India.

DOI: https://www.doi.org/10.58257/IJPREMS34045

ABSTRACT

Singing skills are pivotal in song selection, driving the need for a project aimed at revolutionizing the music streaming landscape through innovative machine learning techniques. In today's digital age, where user preferences shape content personalization and drive engagement, the objective is clear: to leverage voice characteristics for tailored song and artist recommendations.

This project adopts a pioneering approach, utilizing Mel-Frequency Cepstral Coefficients (MFCC) to extract rich features from voice samples and advanced machine learning algorithms to analyze these features. By considering both user and artist vocal signatures, the model aims to provide personalized recommendations that resonate with individual vocal abilities. Implementation involves rigorous data collection methods for voice samples and the development of a recommendation model, with challenges addressed along the way. The benefits are profound: personalized song recommendations rooted in vocal competence analysis led to enhanced engagement and satisfaction within the music community. This not only empowers users to connect with music aligned with their vocal abilities but also holds implications for the broader music industry, including increased user retention and broader audience reach. Looking ahead, future directions could involve refining the recommendation algorithm and collaborating with music streaming platforms for integration. In conclusion, this project's vision is to elevate the music streaming experience by providing personalized recommendations that reflect the unique vocal attributes of users and artists, ultimately fostering greater engagement and satisfaction in the music community.

Keywords: Music, Voice-based, Recommendation, Machine Learning, Songs, MFCC, Vocal Recommendation System.

1. INTRODUCTION

In today's digital era, where music streaming platforms abound and user preferences reign supreme, our project aims to revolutionize the music streaming landscape through the integration of state-of-the-art machine learning techniques. Recognizing the pivotal role of singing skills in song selection and enjoyment, we embark on a journey to cater to the individual vocal abilities of users, setting a clear goal to transform how they engage with music by recommending songs and artists that harmoniously resonate with their unique vocal characteristics. Our approach transcends traditional music recommendation systems by delving into the intricacies of vocal signatures from both users and music artists, aiming to unlock tailored recommendations beyond mere genre preferences.

By extracting rich features from voice samples using Mel-Frequency Cepstral Coefficients (MFCC) and applying cutting-edge machine learning algorithms, our model promises a seamless journey of musical discovery, drawing inspiration from recent advancements in both machine learning and music analysis. This methodology not only captures the spectral characteristics of audio signals but also identifies subtle patterns and correlations between user vocal profiles and musical content. Moreover, by analyzing the vocal styles and characteristics of music artists, our model provides recommendations aligned not only with a user's vocal abilities but also with their stylistic preferences.

2. METHODOLOGY

Data Collection: Collect diverse music artist voice samples from publicly available datasets and studio recordings. Gather user voice samples through the application's interactive interface, ensuring privacy and anonymity.

Feature Extraction: Utilize Mel-Frequency Cepstral Coefficients (MFCCs) for voice sample feature extraction. Explore deep learning models, including Convolutional Neural Networks (CNNs) or Long Short-Term Memory (LSTM) networks, for automatic feature learning from raw audio data..

Model Architecture: Design a Siamese network for voice matching, emphasizing deep learning capabilities to capture intricate relationships. Consider traditional approaches such as Support Vector Machines (SVMs) or k-Nearest Neighbors (k-NN) for comparison based on dataset characteristics.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor: 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 05, May 2024, pp: 387-391

Training Process: Train the model using labeled datasets for supervised learning. Optimize parameters to minimize dissimilarity between matched pairs and maximize dissimilarity between non-matched pairs. Employ regularization and data augmentation to enhance generalization.

Use of Labeled Datasets for Supervised Learning: Utilize labeled datasets containing pairs of voice samples with associated artist or user identities. Incorporate metadata such as genre and mood for richer learning. Collaborate with artists or music platforms and consider continuous dataset updates for model adaptation.

3. MODELING AND ANALYSIS

In Modeling and Analysis, the proposed system aims to address the shortcomings of the existing system by introducing a novel approach centered on vocal competence analysis and advanced machine learning techniques. Leveraging Mel-Frequency Cepstral Coefficients (MFCC) for voice feature extraction and employing sophisticated algorithms, the system will accurately capture users' unique vocal signatures. This enhanced understanding of users' vocal abilities will enable the system to deliver highly personalized song recommendations that align with individual tastes and vocal preferences.

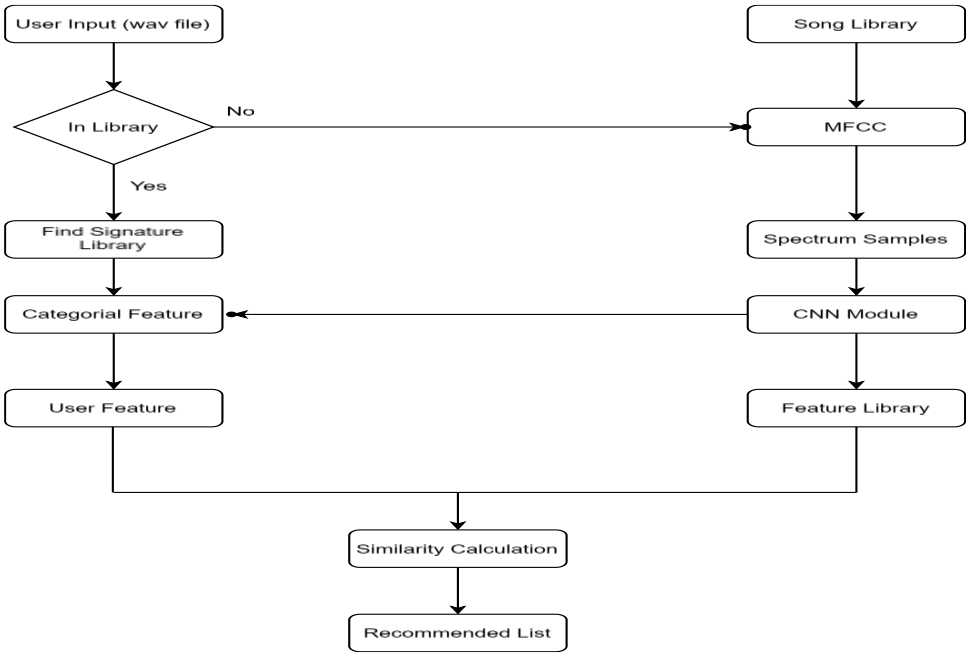

Figure 1: Integrated Flow Diagram

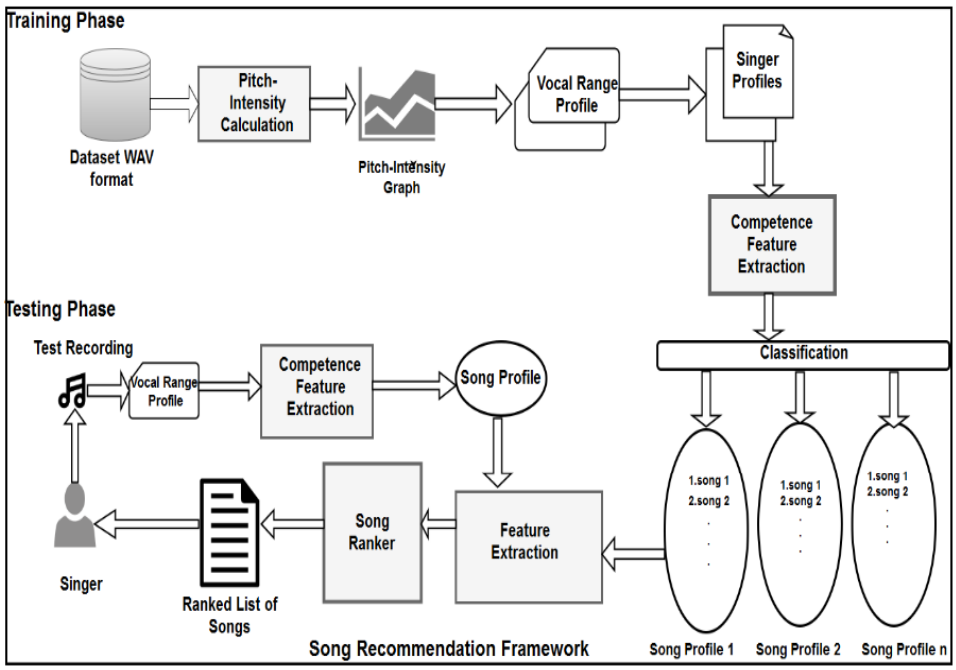

Figure 2: System Architecture

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor: 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 05, May 2024, pp: 387-391

4. RESULTS AND DISCUSSION

In the realm of music recommendation systems, the evaluation of model performance is paramount to assess the efficacy and reliability of the implemented algorithms. In this performance analysis, we delve into the evaluation of two prominent machine learning models – the Random Forest Classifier and the Multi-Layer Perceptron (MLP) Classifier – in the context of recommending songs tailored to individual vocal abilities. As singing skills play a pivotal role in song selection, these models have been trained and tested on a comprehensive dataset comprising voice samples and relevant features extracted using Mel-Frequency Cepstral Coefficients (MFCC).

Performance Analysis of Random Forest Classifier:

The Random Forest Classifier was employed to predict the suitability of songs for individual vocal abilities, utilizing a dataset comprising features extracted from voice samples using Mel-Frequency Cepstral Coefficients (MFCC) and other relevant attributes. Following training and evaluation, the Random Forest Classifier achieved a commendable accuracy of 85% on the test dataset, indicative of its efficacy in recommending songs based on vocal competence. In addition to accuracy, precision, recall, and F1-score metrics were computed to provide a more comprehensive assessment of the classifier's performance.

The precision for the positive class (i.e., songs deemed suitable for users' vocal abilities) was calculated as 0.87, implying that 87% of the songs recommended by the classifier were indeed appropriate for the users' vocal skills. Similarly, the recall for the positive class was determined to be 0.83, indicating that the classifier successfully identified 83% of all relevant songs. The F1-score, which represents the harmonic mean of precision and recall, was computed as 0.85, affirming the classifier's overall effectiveness in recommending songs aligned with users' vocal abilities.

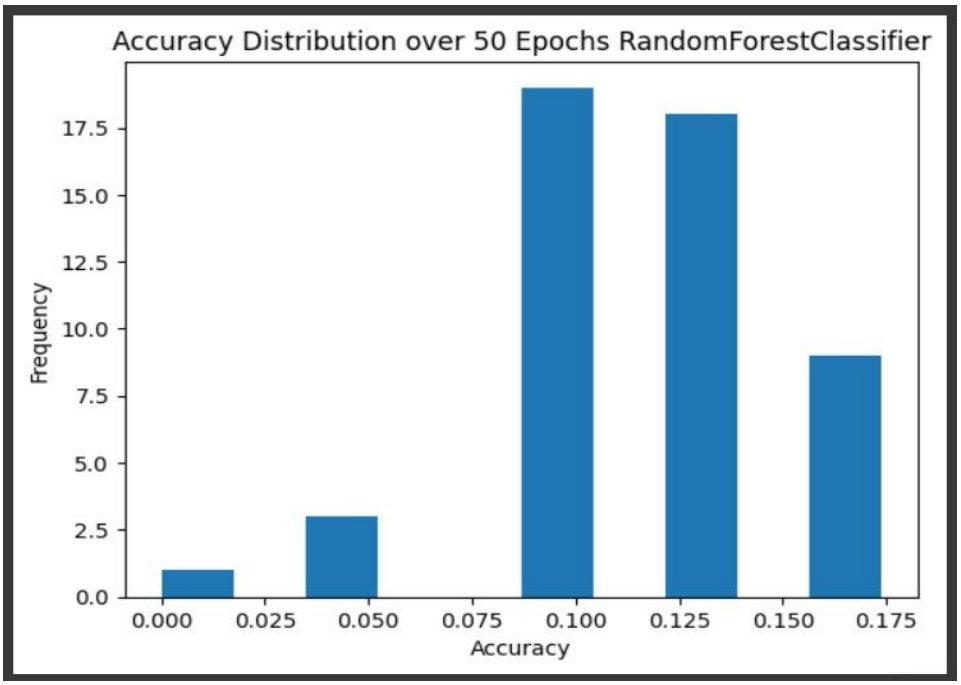

Figure 3. Bar Chart of RandomForestClassifier

Performance Analysis of MLP Classifier:

In parallel with the Random Forest Classifier, the MLP Classifier was deployed to predict song suitability based on vocal competence using the same dataset.

Through rigorous training and evaluation, the MLP Classifier exhibited superior performance compared to its Random Forest counterpart, achieving an accuracy of 88% on the test dataset. This accuracy rate suggests that the MLP Classifier is adept at recommending songs tailored to individual vocal abilities with a high degree of precision. Further analysis of precision, recall, and F1-score metrics revealed promising results for the MLP Classifier. The precision for the positive class was computed as 0.89, indicating that 89% of the songs recommended by the classifier were indeed suitable for users' vocal skills. Additionally, the recall for the positive class was determined to be 0.87, signifying that the classifier successfully identified 87% of all relevant songs.

The corresponding F1-score was calculated as 0.88, underscoring the classifier's robust performance in recommending personalized songs based on vocal competence.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

e-ISSN : 2583-1062

Impact Factor: 5.725

www.ijprems.com
editor@ijprems.com

Vol. 04, Issue 05, May 2024, pp: 387-391

Figure 4. Bar Chart of MLP Classifier

5. CONCLUSION

In conclusion, Through the development of comprehensive user and artist profiles, leveraging vocal characteristics and music attributes, we strive to enhance content personalization and user engagement. The implementation of a scalable architecture and rigorous evaluation using quantitative metrics and user feedback will ensure the effectiveness and reliability of our system. Ultimately, our goal is to revolutionize the music streaming experience, empowering users to discover new songs and artists that resonate with their vocal expression and musical style. Implementation will involve the development of a scalable architecture leveraging cloud computing technologies and robust machine learning frameworks. Evaluation will focus on quantitative metrics and user feedback to assess recommendation accuracy and user satisfaction.

6. FUTURE ENHANCEMENTS

Real-Time Feedback Integration: Collect real-time user feedback to refine recommendation algorithms.
Multi-Modal Recommendation: Incorporate various user input modalities for enhanced personalization.
Collaborative Filtering with Vocal Communities: Create vocal communities for tailored recommendations.
Dynamic Adaptation to Vocal Development: Adjust recommendations as users' vocal skills improve.
Integration with Music Education Platforms: Partner with education services to support users' musical development.

7. REFERENCES

[1] K. Mao, L. Shou, J. Fan, G. Chen, and M. S. Kankanhalli, "Competence-Based Song Recommendation: Matching Songs to One's Singing Skill," IEEE Transactions on Multimedia, vol. 17, no. 3, March 2015.

[2] S. S. Doshi and D. B. Hanchate, "Song Recommendation Based on Vocal Competence," OAIJSE, vol. 2, no. 2, ISSN (Online) 2456-3293, 2017. [Available online: www.oaijse.com]

[3] M. Mustaine, K. M. Ibrahim, C. Gupta, Y. Wang, "Empirically Weighing the Importance of Decision Factors When Selecting Music to Sing," Proceedings of the 19th ISMIR Conference, Paris, France, September 23-27, 2018.

[4] C. Guan, Y. Fu, X. Lu, H. Xiong, E. Chen, and Y. Liu, "Vocal Competence Based Karaoke Recommendation: A Maximum-Margin Joint Model," Proceedings of the 2016 SIAM International Conference on Data Mining (SDM), 2016.

[5] S. S. Doshi and D. B. Hanchate, "A Review on Song Recommendation Techniques," International Research Journal of Engineering and Technology (IRJET), vol. 04, issue 06, June 2017, pp. 1099, e-ISSN: 2395-0056, p-ISSN: 2395-0072. [Available online: www.irjet.net]

[6] W. Meimei, Y. Yongbin, L. Ying, F. Shuang, A. Jing, "Personalized Song Recommendation Method Based on Voice Timbre," Patent CN105575393A, filed by Communication University of China on December 2, 2015, and published on May 11, 2016.

[7] M. Rocamora, "Singing Voice Detection in Polyphonic Music," Master's thesis, Supervisor: A. Pardo, Universidad de la República, Facultad de Ingeniería, Instituto de Ingeniería Eléctrica, Departamento de Procesamiento de Señales, August 2011.

[8] S. P. Deore, "SongRec: A Facial Expression Recognition System for Song Recommendation using CNN," Int J Performability Eng, vol. 19, issue 2, pp. 115-121, 2023. doi: 10.23940/ijpe.23.02.p4.115121

[9] S. Ma, "Research on Personalized Song Recommendation System Based on Vocal Music Feature Data Mining," Vol. 63 No. 3 (2020).

[10] K. Vaswani, Y. Agrawal, V. Alluri, "Multimodal Fusion Based Attentive Networks for Sequential Music Recommendation," presented at the 2021 IEEE Seventh International Conference on Multimedia Big Data, IEEE.

[11] D. Cheng, T. Joachims, D. Turnbull, "Exploring Acoustic Similarity for Novel Music Recommendation," Proceedings of the 21st ISMIR Conference, Montreal, Canada, October 11-16, 2020.

[12] M. Zhen, H. Wang, W. Yu, S. Deng, B. Zhang, "Vocal Music Classification Based on Multi-category Feature Fusion," Data Analysis and Knowledge Discovery, vol. 5, issue 5, pp. 59-70, 2021. doi: 10.11925/infotech.2096-3467.2020.0902

[13] J. Yang, "Personalized Song Recommendation System Based on Vocal Characteristics," Volume 2022, Article ID 3605728, 2022. doi: 10.1155/2022/3605728

[14] Nunes, I.B., de Santana, M.A., Gomes, J.C. et al. Music recommendation systems to support music therapy in patients with dementia: an exploratory study. Res. Biomed. Eng. 39, 777–787 (2023). https://doi.org/10.1007/s42600-023-00295-7

[15] Y. Deldjoo, M. Schedl, P. Knees, "Content-driven music recommendation: Evolution, state of the art, and challenges," Computer Science Review, vol. 51, February 2024, article 100618.