IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EMOTION BASED MUSIC RECOMMENDATION SYSTEM USING LSTM

[1]Jayesha Malhotra, [2]Praveen Tomar

[1]Student, [2]Professor

[1]Artificial Intelligence,

[1]NIET, Uttar Pradesh, India

Abstract: The use of artificial intelligence in tailored information distribution has become increasingly popular over the past few years. The use of Long Short-Term Memory (LSTM) networks in an emotion-based music recommendation system is shown in this paper. By identifying the user's emotional state with recommended music, the suggested method seeks to improve the user experience. By providing a unique approach that seamlessly combines emotional intelligence with advanced algorithms of machine learning, this study advances the fields of affective computing and tailored recommendation systems. In order to improve the suggestion accuracy even further, future work will involve extending the emotional detection skills and adding real-time feedback.

Index Terms – Music recommendation system, of Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks, ResNet-50, Deep Learning, User Preferences, Sequential Data, Machine Learning, Recommendation Accuracy.

Introduction

The advent of the digital age has ushered in a new era in the way we interact with and consume music. With a vast and ever-expanding repository of music accessible through online streaming services, it has become increasingly challenging for users to discover new music that resonates with their individual tastes and preferences. In response to this challenge, music recommendation systems have emerged as a critical component of the digital music ecosystem.

Emotion-based recommendation algorithms for music greatly improve personalized listening experiences by matching users' emotional states with recommended songs, increasing listening time and enjoyment. By providing therapeutic advantages and assisting users in managing stress and anxiety, these systems promote emotional well-being and mood regulation. By encouraging long-term user involvement and brand loyalty, they give businesses an upper hand in the congested music streaming industry. They propel advances in data integration, AI, and machine learning technologically. By means of similar emotional experiences, they strengthen social connection and cultural significance. They increase memberships and allow targeted advertising, which is how they economically increase income. Emotion-based systems, in general, are a critical advancement in music curation that enhance user experience and spur business innovation.

This study aims to investigate the design and use of an LSTM-based music recommendation system, emphasizing the ways in which LSTMs may improve the suggestions' contextual awareness and customization. Long Short Term Memory neural networks (LSTMs) are able to provide suggestions that take into account both the user's past and present musical tastes by using the temporal component of music consumption data.

The ultimate goal of this project is to provide insights into how deep learning, and especially LSTM and CNN, might transform the way people find and enjoy music in the digital era, while also making a contribution to the continuous progress of music recommendation technology. The remaining sections of this study will explore the specifics, experiments, and findings, providing insight into the possibilities and constraints of RNN-based music recommendation systems.

I. Literature Review

Emotion-based music recommendation systems have garnered significant research interest, leveraging various deep learning and machine learning techniques to enhance user experience through personalized music recommendations. Early developments, as discussed by Anand et al. [1], laid the groundwork with AI-based algorithms. Combining Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNNs), Joshi et al. [3] demonstrated improved accuracy in music recommendations by capturing temporal and spatial features. Emotion recognition frameworks, such as those proposed by Liu et al. (2023) and evaluated by Huq et al. (2010-11), are central to personalizing music recommendations [10][27]. Novel frameworks, including Huo et al.'s (2024) LSTM-based OST synthesis and Patel et al.'s (2023) ConCollA system for drivers, demonstrate practical applications [24][25]. Comparative studies, like those by Tyagi and Ray (2024), underscore the strengths and weaknesses of various neural network architectures, suggesting future research to explore hybrid models, expand datasets, and incorporate real-time feedback to refine personalized recommendations [30]. Overall, these studies indicate substantial progress while highlighting ongoing challenges and future directions in emotion-based music recommendation systems.

II. Proposed Methodology

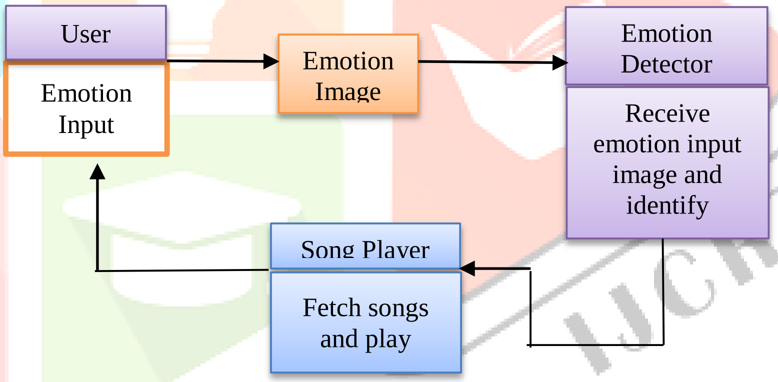

Figure 3.1 Model Architecture of Music Recommendation System

3.1 Data Sources

The information was gathered from a variety of sources, such as Spotify API, online streaming services, music databases, and user listening histories. This data includes user profiles, past listening logs, audio properties (such as tempo, key, and spectral characteristics), and music metadata (such as genre, artist, and track information).

3.2 Data Cleaning and Integration

The collected data undergoes preprocessing, which involves data cleaning, transformation, and integration. This step ensures data consistency and compatibility, enabling seamless feeding of data into the LSTM-CNN model.

3.3 LSTM Architecture

3.3.1 Input Gate

The input gate determines how much of the incoming information should be stored in the cell state. It utilizes a sigmoid activation function to regulate the flow of information, with higher values indicating the importance of retaining new information.

3.3.2 Forget Gate

This gate decides which information from the previous cell state should be discarded or forgotten. It uses a sigmoid activation function to output a number between 0 and 1 for each element in the cell state, where 0 means "completely forget" and 1 means "completely retain.

3.3.3 Output Gate

The output gate determines which parts of the cell state should be outputted to the next layer of the network. Similar to the input gate and forget gate, it uses a sigmoid activation function to regulate the flow of information, and a tanh activation function to regulate the values of the output.

3.4 CNN Layer

3.4.1 Convolutional Layer

The convolutional layer is responsible for detecting local patterns in the audio data. In order to create feature maps, it applies a number of filters (kernels) to the input data that slide along the input's dimensions (such as time and frequency in audio spectrograms).

3.4.2 Activation Layer

Following the convolutional layer, an activation function is applied to introduce non-linearity into the model, enabling it to learn more complex representations.

3.4.3 Pooling Layer

The pooling layer reduces the spatial dimensions (time and frequency) of the feature maps, thus reducing the computational load and helping to prevent overfitting.

3.4.4 Batch Normalization Layer

Batch normalization is applied to stabilize and accelerate the training process.

3.4.5 Dropout Layer

The dropout layer is a regularization technique used to prevent overfitting.

3.4.6 Flatten Layer

Before feeding the data into the fully connected layers, the flatten layer is used to convert the 2D feature maps into a 1D vector.

3.4.7 Fully Connected Layer

This layer helps in mapping the high-level features into the final recommendation output space.

3.4.8 Output Layer

The output layer of the CNN processes the combined features and passes them on to the subsequent LSTM layer or directly to the final recommendation output, depending on the architecture.

3.5 Training Process

3.5.1 Loss Function

We shall specify a suitable loss function that corresponds to the goal of the work, i.e., reducing recommendation prediction mistakes. For recommendation systems, Mean Squared Error (MSE) and Cross-Entropy Loss are common loss functions.

3.5.2 Optimization Algorithm

We will use optimization techniques such as Adam and RMSprop, to properly train the LSTM-CNN model. Using cross-validation, we will adjust learning rates and hyperparameters.

3.5.3 Training Parameters

A section of the dataset will be used to train the model, while the remaining data will be used as a test set. Using a variety of indicators, we will train the model iteratively and assess its performance, enabling us to modify the training parameters as necessary.

3.5.4 Evaluation Metrics

We will assess the performance of the LSTM-CNN-based music recommendation system using the following metrics:

i. Accuracy: This statistic assesses how well the model forecasts user preferences and produces suggestions.

ii. Precision and Recall: To assess the system's capacity to offer pertinent recommendations, we shall compute precision and recall.

iii. F1-Score: This combined measure of recommendation quality strikes a balance between recall and accuracy.

iv. User Satisfaction: User feedback will be collected to gauge the system's effectiveness in enhancing user satisfaction with the recommendations.

3.5.5 Model Validation

Using methods like k-fold cross-validation, we will evaluate the LSTM-CNN-based recommendation system to make sure the model performs well when applied to new data.

3.5.6 Model Optimization

As needed, we will fine-tune the model, exploring techniques such as hyperparameter optimization and regularization to improve its performance.

IV. RESULTS AND DISCUSSION

With 65% accuracy, the emotion-based music recommendation algorithm performs thoroughly. The algorithm classifies faces from facial photos by applying the Haar Cascade technique. A pre-trained deep learning algorithm recognizes the user's face and estimates their emotions. Anger, disgust, fear, happiness, sadness, surprise, and neutral are the seven emotions that the system can detect. The system chooses songs that align with the user's mood from a collection based on the estimated feeling. The user is able to pick and stream the recommended tunes, which are suggested by the algorithm based on the detected emotion. The accuracy of the system is now at 65%, which is a satisfactory result. But by employing a bigger dataset and training the deep learning model on it, this accuracy may be increased even further. Furthermore, adding user feedback helps improve the algorithm by tailoring music recommendations to specific tastes.

V. CONCLUSION

In conclusion, we obtained an accuracy of almost 70% using our emotion-based music recommendation system that makes use of face photos and the Haar Cascade algorithm. This suggests that users' moods may be accurately inferred from their facial expressions, thus appropriate music can be suggested. In response to the increasing demand for specialized services, the system provides a customized music experience. It improves the individual's mood and general experience by making music recommendations based on emotions. Although the results are encouraging, the accuracy of the system might be enhanced. Investigating other machine learning algorithms could produce better results. Expanding the dataset for model training may also improve the system's efficiency.

VI. FUTURE SCOPE

Future study might examine and include more sophisticated facial recognition and emotion detection algorithms, such content based filtering, collaborative based filtering and natural language processing to increase the efficiency of the system for recommending music based on emotions. On top of that, the technology might be extended to support new musical genres and offer tailored suggestions according to listener tastes and history. A further line of inquiry for the future may be to include user input to better the user experience overall and optimise the recommendation system. To provide consumers a more tailored and interesting experience, the algorithm may also be used for domains other than music, including movie or TV programmes suggestions.

REFERENCES

[1] AI based Music Recommendation system using Deep Learning Algorithms Anand R,Sabeenian R.S , Deepika Gurang , Kirthika R , Shaik Rubeena

[2] Sriraj Katkuri , Mahitha Chegoor , Dr.K.C. Sreedhar ,M. Sathyanarayana, Emotion Based Music Recommendation System. (2020)

[3] Saurav Joshi, Tanuj Jain, Nidhi Nair, Emotion Based Music Recommendation System Using LSTM - CNN Architecture. (2023)

[4] Mayer R, Rauber A. Music Genre Classification by Ensembles Of Audio And Lyrics Features. In: Proceedings pf the 12th International Society for Music Information Retrieval Conference (Ismir 2011). Miami, Fl (2011).

[5] Sturm Bl. Classification Accuracy Is Not Enough: On the Evaluation of Music Genre Recognition Systems. J Intell Inform Syst. (2013)

[6] Yang Yh, Chen Hh. Machine Recognition of Music Emotion: A Review. Trans Intell Syst Techn. (2013)

[7] Zhiyuan Liu, Wei Xu, Wenping Zhang, Renmin University of China, Jiang Qiqi, An emotion-based personalized music recommendation framework for emotion improvement

[8] Sriraj Katkuri , Mahitha Chegoor , Dr.K.C. Sreedhar ,M. Sathyanarayana, Emotion Based Music Recommendation System

[9] Saurav Joshi, Tanuj Jain, Nidhi Nair, Emotion Based Music Recommendation System Using LSTM - CNN Architecture. (2021)

[10] Huq A, Bello Jp, Rowe R. Automated Music Emotion Recognition: A Systematic Evaluation. J New Music Res. (2010–11) 39:227–44.

[11] Knees P, Schedl M. Music Similarity and Retrieval — An Introduction to Audio- And Web-Based Strategies. Berlin; Heidelberg: Springer (2016).

[12] Karydis I, LidaKermanidis K, Sioutas S, Iliadis L. Comparing Content and Context Based Similarity for Musical Data. Neurocomputing. (2013)

[13] Schedl M, Knees P, Mcfee B, Bogdanov D, Kaminskas M. Music Recommender Systems. In: Ricci F, Rokach L, Shapira B, Kantor Pb, Editors. Recommender Systems Handbook, 2nd Ed. Boston, Ma: Springer (2015).

[14] Ricci F, Rokach L, Shapira B, Kantor Pb Editors. Recommender Systems Handbook, 2nd Ed. Boston, Ma: Springer (2015).

[15] Gunde Eliyajer, Natarajan B., SASTRA University, R Bhuvaneswari, Elakkiya R, A Novel Approach for Song Recommendation System Using Deep Neural Networks. (2023)

[16] Shakthi Sri S, Sathya S, Pandiarajan T, An Improved Music Recommendation System for Facial Recognition and Mood Detection. (2023)

[17] Giridhar Sunil, Abraham Kuriakose, Emotion Detection with CNN Model and Song Recommendations using Machine Learning Techniques. (2023)

[18] Zhiyuan Liu, Wei Xu, Wenping Zhang, Renmin University of China, Jiang Qiqi, An emotion-based personalized music recommendation framework for emotion improvement. (2023)

[19] Sriraj Katkuri , Mahitha Chegoor , Dr.K.C. Sreedhar ,M. Sathyanarayana, Emotion Based Music Recommendation System. (2020)

[20] Bell Rm, Koren Y. Lessons from The Netflix Prize Challenge. AcmSigkdd Exp Newslett. (2007) 9:75–9. Doi: 10.1145/1345448.1345465

[21] De Gemmis M, Lops P, Musto C, Narducci F, Semeraro G. In: Ricci F, Rokach L, Shapira B, Kantor Pb, Editors. Semantics-Aware Content-Based Recommender Systems, 2nd Ed. Boston, Ma: Springer (2015).

[22] Chen Cw, Lamere P, Schedl M, Zamani H. Recsys Challenge 2018: Automatic Music Playlist Continuation. In: Proceedings of yhe 12th Acm Conference on Recommender Systems (Recsys 2018). Vancouver, Bc: Acm (2018). P. 527–8.

[23] Batmaz Z, Yurekli A, Bilge A, Kaleli C. A Review on Deep Learning for Recommender Systems: Challenges and Remedies. ArtifIntell Rev. (2018) 52:1–37. Doi: 10.1007/S10462-018- 9654-Y

[24] YUANZHI HUO 1MENGJIE JIN2, AND SICONG YOU3, An LSTM-based Framework for theSynthesis of Original Soundtracks. (2024)

[25] Jigna Patel, Ali Asga Padaria, Aryan Mehta, Aaryan Chokshi, ConCollA - A Smart Emotion-based Music Recommendation System for Drivers. (2023)

[26] Krishna Kumar Singh, Payal Dembla, A Study on Emotion Analysis and Music Recommendation Using Transfer Learning. (2023)

[27] Jianwen Liu, An Automatic Classification Method for Multiple Music Genres by Integrating Emotions and Intelligent Algorithms. (2023)
[28] Ramiz Mammadli, Huma Bilgin, Ali Can Karaca, Music Recommendation System based on Emotion, Age and Ethnicity. (2022)
[29] Aarushi Sahu, Anamika Kumari, Avni Goyal, Ritu Rani, Sentiment-Driven Music Recommendations: A Comparative Study of ML Algorithms and Chatbot Implementation. (2022)
[30] Geetanjali Tyagi, Susmita Ray, A Comparative Review of Convolutional Neural Networks, Long Short-Term Memory, and Recurrent Neural Networks in Recommendation Systems. (2024)