

# Song Recommendation based on user's Activity using Ensemble Learning and Clustering

*Deepali Joshi, Akshay Gade, Phalguni Savale, Vinay Bhujbal, Pranavraj Goje and Saniya Mhamane .*

Department of Artificial Intelligence and data science, Vishwakarma Institute of Technology, Pune, India, 411037.

**Abstract.** The Song Recommendation System Based on User Schedule project is designed to provide users with personalized music recommendations that match their daily activities and mood swings. With a busy and hectic schedule, it can be challenging to find time to select music that matches a user's current activity and mood. This project aims to provide a solution to this problem by analyzing the user's daily schedule, including their planned activities and time of day, and using machine learning algorithms to recommend songs that fit their mood and energy level during each activity. The project utilizes a variety of technologies, such as React.js for the front-end and various machine learning algorithms using python for the back-end, to provide a user-friendly interface that allows users to input their schedules and receive song recommendations.

**Keywords.** Ensemble learning, XGBoost, Firebase, K-means, Stacking approach, Clustering

## 1. Introduction

About The Song Recommendation System Based on User Schedule project is a cutting-edge solution that provides users with personalized music recommendations based on their daily activities and mood. The system uses machine learning algorithms to analyze the user's daily schedule and recommend songs that match their mood and energy level during each activity. The system calculates the user's mood based on the intensity, length, and heart rate of the activity before utilizing the K-means algorithm to group the activities according to mood.

The system is designed to provide personalized song recommendations for different activities like working, exercising, relaxing, or studying. For example, if the user is working for a long time and the heart rate is high, the mood can be classified as stressed, and the system will recommend songs that match the user's mood, like instrumental or calming music. Similarly, if the user is doing a high-intensity workout, the mood can be classified as energetic, and the system will recommend songs that match the user's mood, like upbeat and fast-paced music.

To implement the recommendation system, the project uses a backend that contains a classification by **stacking approach (ensemble learning) of XGboost, Random forest, K-nearest neighbor algorithm** and clustering using **K-means algorithm**. The classification model has a **93.69%** accuracy rate, which makes it a very trustworthy answer. The front-end of the system is developed using React.js, which provides a user-friendly interface that allows users to input their schedules and receive song recommendations. The system's backend is developed using node.js, which is used to connect and communicate with the Firebase real-time database. This allows for the efficient handling of data on the server-side, which is crucial for providing quick and accurate recommendations to the user. Additionally, **Node.js** has a vast ecosystem of modules and libraries, which can be leveraged to further enhance the functionality and performance of the project.

In this song recommendation project, Firebase is used as a real-time database to store and manage user data. A NoSQL cloud-hosted database, the Firebase real-time database enables users to store and sync data in real-time across several clients. In this project, Firebase is used to store user information including listening preferences and listening history and to update suggestions in real-time depending on changes to the user's information.

## 2. Literature Review

[1] The research suggests a theory-driven method for suggesting upbeat music for daily tasks. Brunel Music Rating Inventory can detect upbeat aspects in music by analyzing audio signals. PepMusic, a song recommender, can properly categorize frequently listened songs for 14 typical daily activities into three high-level latent activity categories, according to the findings of a preliminary user evaluation. This has significance for suggesting music for everyday activities.

The paper [2] describes a method for determining the mood of a song by using a Multiclass Neural Network for Classification. The network recognizes emotional and mood-related characteristics in music and categorizes the song into a specific mood category. The goal is to automate the process of assigning a mood to a song, which can be useful for music recommendation systems, film scoring, and other applications. The machine learning algorithm used is a multi-class neural network classifier that can handle multiple output classes and is well-suited for this task.

[3] The authors have developed a mood-based music player that uses real-time mood detection from analyzing the user's facial expressions to suggest songs that reflect the user's present mood. The system uses a CNN algorithm and a MobileNet model with Keras to detect and classify various human emotions. The aim is to provide personalized music recommendations and increase customer satisfaction.

Personalized Song Recommendation System Based on Vocal Characteristics [4] the authors of the study provide a technique for song recommendations that makes advantage of voice cues in order to enhance user experience. CNN is used with a threshold model to increase the categorization accuracy of songs.

Music Recommendation System Based on User's Sentiments Extracted from Social Networks [5] proposes the improved Sentiment Metric (eSM), which is based on a lexicon-based sentiment metric, is used in the research to develop a music recommendation system. Based on the current emotion intensity, the algorithm recommends music by extracting feelings from social media. The project uses a lexical-based approach to capture the emotional content of music, which is then matched with the sentiment intensity in social media to provide appropriate music recommendations. The study suggests that eSM can be used to generate personalized music recommendations for users based on their current emotional state.

[6] The paper reviews speech-based emotion recognition in human-computer interaction by using deep learning techniques. It describes how DNN and CNN are used for emotion extraction from speech signals, and how deep learning methods can compute various non-linear components for emotion recognition and also the paper offers a research of speech-based emotion identification using deep learning algorithms.

[7] presents a chatbot-based music recommendation system that recommends songs on the basis of the user's text tone. The system uses IBM Tone Analyzer API and Last.FM API to identify the emotion and recommend songs respectively. The chatbot uses Support Vector Machine (SVM) with One-vs-Rest paradigm to predict the tone of new texts. The system divides the songs based on their energy and stress levels to suggest songs that match the user's mood. The final tone is determined by identifying those predicted with at least a 0.5 probability.

The paper [8] proposes an approach to improve the accuracy of emotion classification from EEG brain signals using feature optimization and the XGBoost algorithm. The technique entails the extraction of features from EEG signals and their optimisation using a procedure that makes use of correlation matrices, information gain, and recursive feature removal. Next, classification is performed using XGBoost. The suggested method was evaluated using the DEAP dataset and evaluated against Naive Bayes, KNN, C4.5, Decision Tree, and Random Forest algorithms as well as other classification techniques.

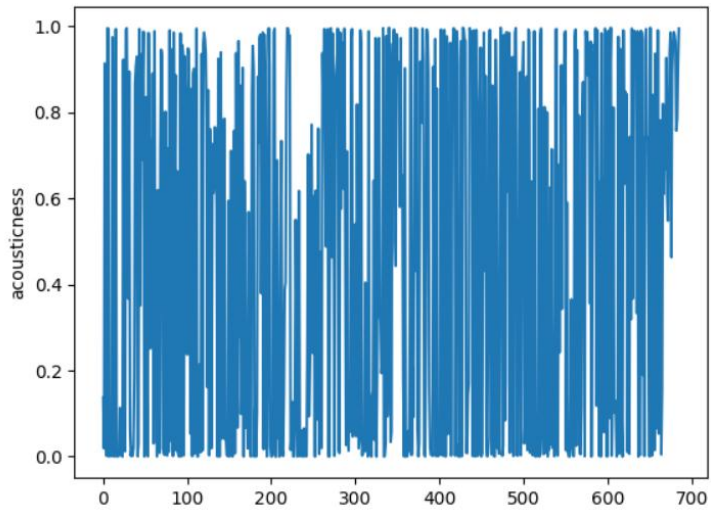
[9] The paper presents a technique for detecting emotions from EEG signals. The technique comprises extracting discriminative features with a Deep Normalized Attention-based Residual Convolutional Neural Network (DNA-RCNN) and classifying data with a modified random forest (M-RF). The proposed method explores attention modules to extract alluring features, which leads to consistent performance. The M-RF algorithm uses an empirical loss function to learn weights on the data subset and assist in precise classification.

[10] This paper reviews current technologies for detecting human emotions, exploring the different sources from which emotions can be read and the technologies developed to recognize them. In addition to reviewing current emotion detection technologies, the paper explores the various domains in which these technologies have been applied, such as affective computing, mental health, and human-computer interaction. The strengths and limitations of existing technologies are also discussed in order to identify areas for further research and improvement.

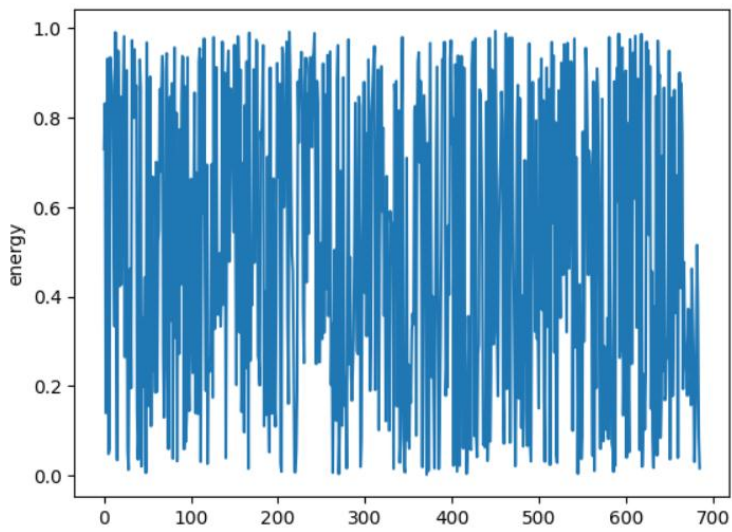
### **3. Methodology**

#### **3.1 Data set**

A supervised dataset about songs with various features has been utilized while designing the proposed system. The dataset includes features like energy, duration, valence, liveliness and instrumentality. Additionally, the dataset also includes a column called mood, which is the target variable. The mood column contains a categorical label indicating the emotional content of each song, with possible values of "happy", "calm", "sad", and "energetic". This makes the dataset suitable for training machine learning models to predict the mood of a song based on its features.



**Fig. 1.** Acousticness



**Fig. 2.** Energy

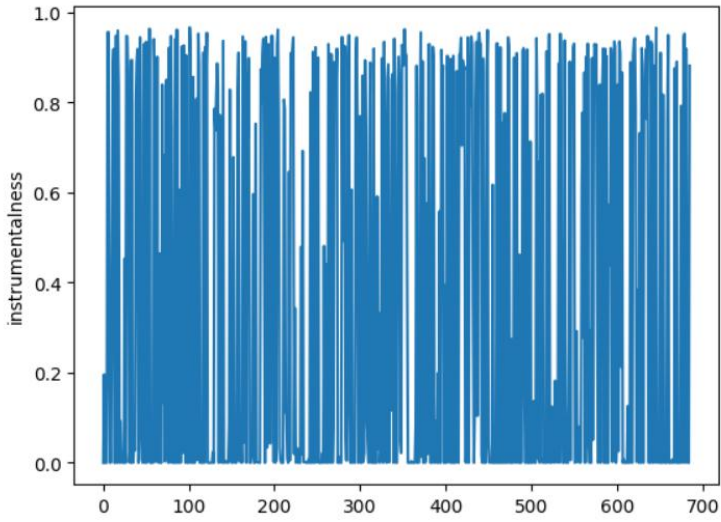


Fig. 3. Instrumentalness

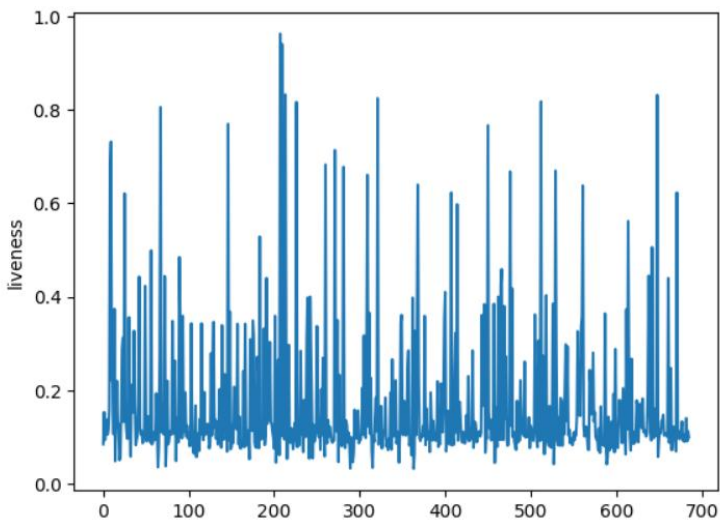


Fig. 4. Liveness

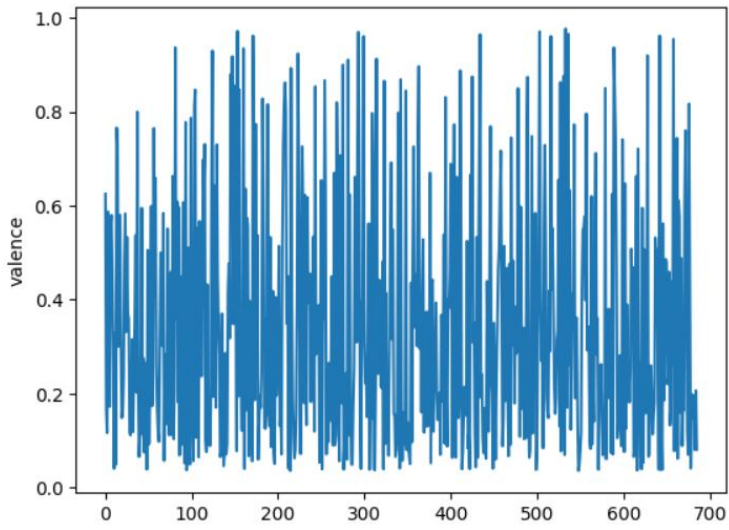


Fig. 5. Valence

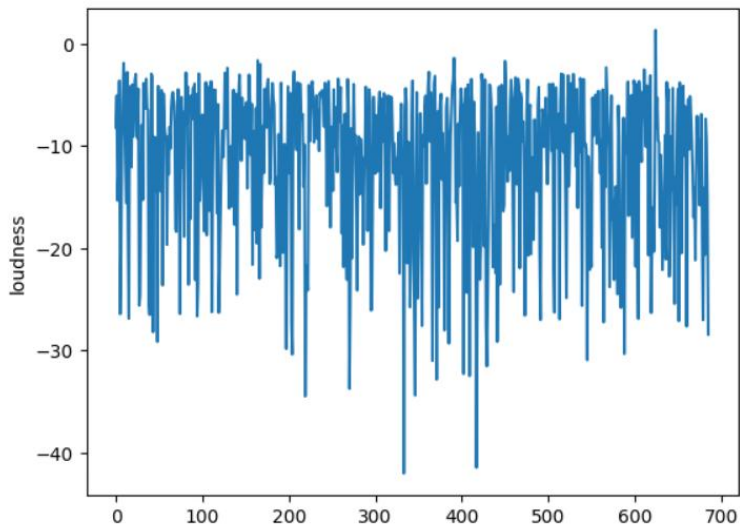


Fig. 6. Loudness

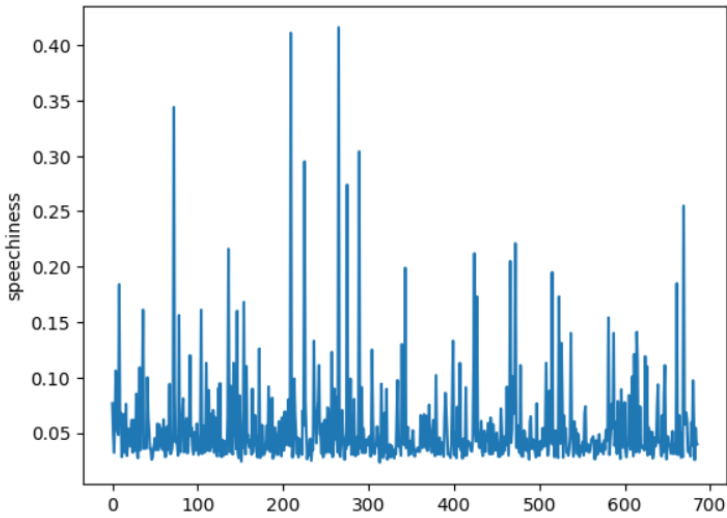


Fig. 7. Speechiness

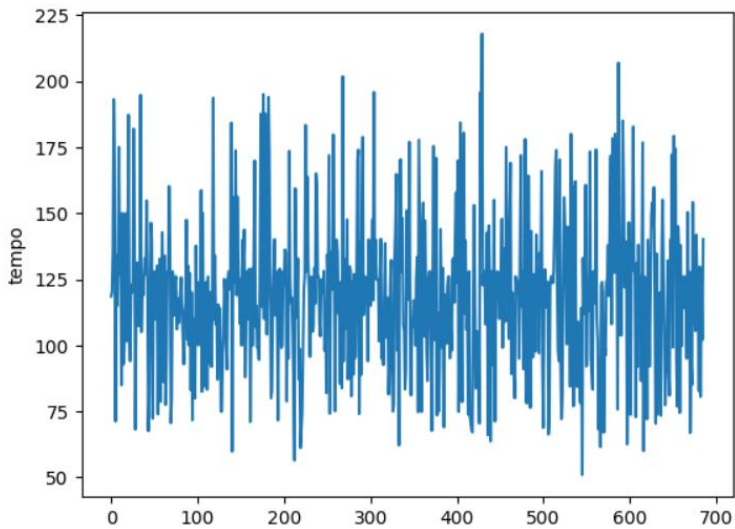


Fig. 8. Tempo

### 3.2 Proposed system

The main goal of this system is to recommend songs based on the mood detected from user activity. For providing this system, the first and foremost goal was to develop a dataset which would contain songs based on mood. In order to develop a dataset with improved classification accuracy, the project utilized several classification algorithms including **XGBoost**, **K Nearest Neighbor**, and **Random Forest**. To further enhance classification accuracy, an **ensemble learning approach** was employed. This involved running all three models to generate predictions, and then selecting the best prediction using a fourth classifier. By combining the strengths of multiple models, the resulting dataset was able to achieve higher levels of accuracy.

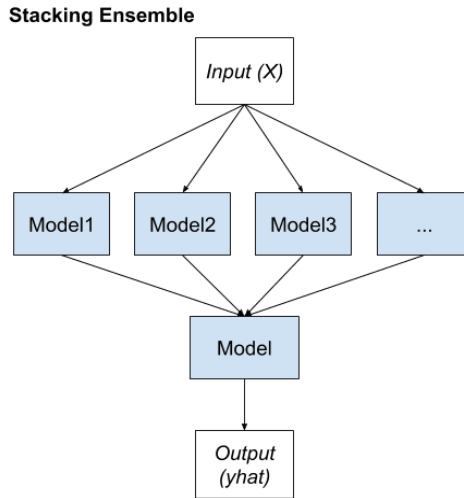


Fig. 9. Stacking Approach

Using this classification model, the data was classified into four major labels - “**Energetic**”, “**Happy**”, “**Calm**”, “**Sad**”.

Second major goal was to detect mood from activity of the user. To achieve this, clustering was performed on a dataset including **heart rate** of users collected from **smartwatch** and a dataset that had activities detected from heart rate. The smartwatch data contained entries of heart rate, gyroscope and gravity values of each and every microsecond of week out of which only heart rate was required therefore the data was trimmed and only heart rate data for each microsecond was obtained. The other data contained activity detected for durations of day, this data was mapped with heart rate and then the final data was obtained according to the requirement.

Using this final data, four clusters were created each one for a mood respectively. Based on the heart rate in each cluster one mood out of the four was assigned to the cluster.

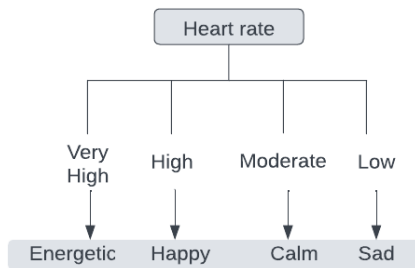


Fig. 10. Mood Clustering

After mapping of cluster and mood the recommendation system became capable of recommending songs to the user from time to time throughout his entire schedule.



### 3.3 Architecture Diagram

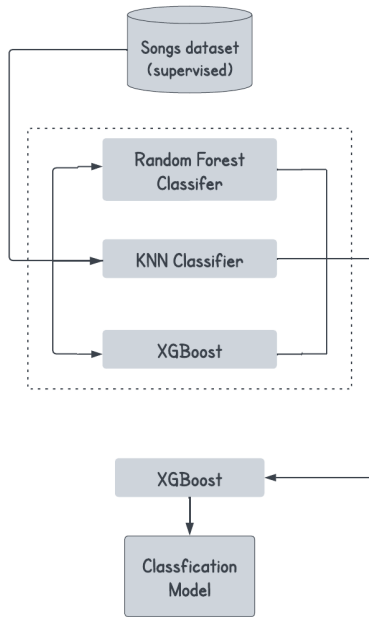


Fig. 11. Classification

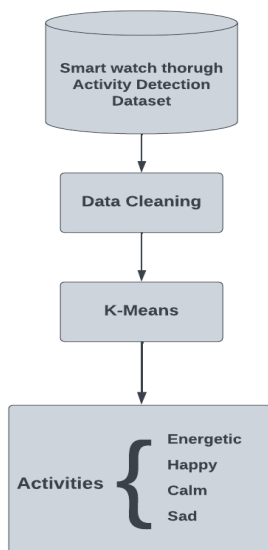


Fig. 12. Clustering

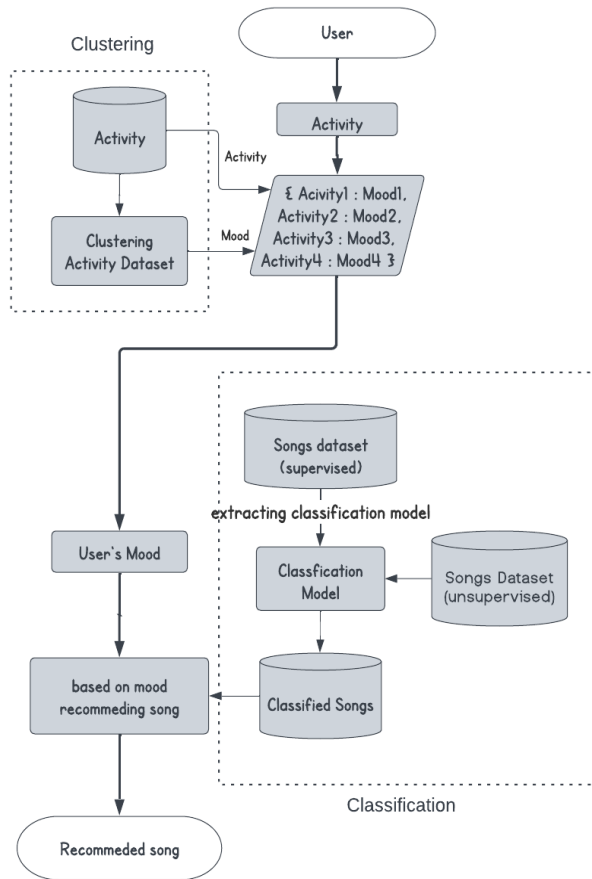


Fig. 13. Proposed system's Flowchart

### 3.3 Algorithm

Step 1: Get the user activity with schedule

Retrieve the user's activity schedule, including details such as the type of activity, duration, and location.

Step 2: Detect mood from user activity using clustered dataset

Use a pre-trained machine learning model to classify the user's activity into a mood cluster (e.g., happy, sad, relaxed, energetic). The model should be trained on a diverse and representative dataset to accurately classify the user's mood.

Step 3: Fetch songs from the classified dataset based on mood detected

Retrieve a list of songs from a classified dataset that are associated with the mood cluster identified in step 2. The classified dataset should include a diverse selection of songs that are associated with different moods, genres, and cultures.

Step 4: Recommend the songs fetched to users from time to time

Recommend the retrieved songs to the user at appropriate times based on their activity schedule. For example, if the user is scheduled to exercise, recommend energetic songs to motivate them. Use a recommender system to personalize the song recommendations based on the user's past listening behavior and feedback.

Step 5: Keep repeating step 2 for each activity present in the user schedule

Continuously monitor the user's activity schedule and update the mood classification based on any changes. For example, if the user changes their schedule to include a yoga session, reclassify their mood as relaxed and fetch appropriate songs from the classified dataset.\

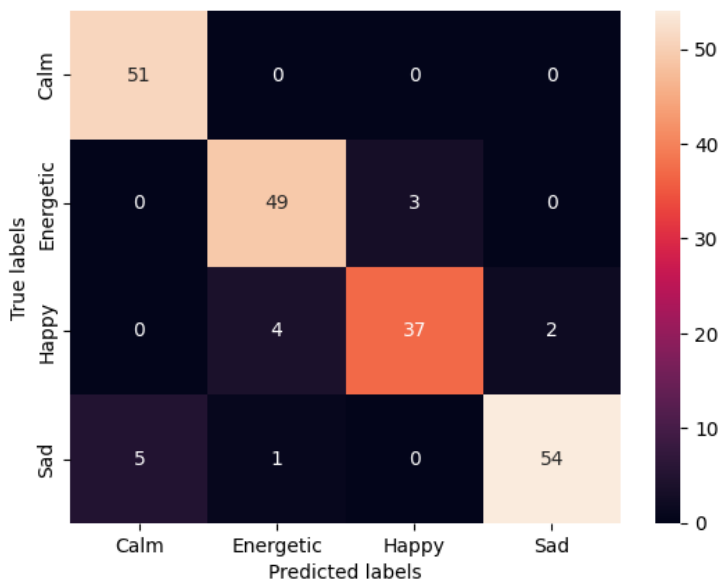
Step 6: Implement a feedback loop

Allow users to rate the recommended songs and provide feedback on their preferences. Use this feedback to refine the recommender system and improve the accuracy of the mood classification.

### 4. Results and Discussions

In evaluating a classification model's performance, confusion matrices are often used. Using the predicted labels compared with the true labels for a set of data, it summarizes how well the model predicted. A confusion matrix normally contains four terms, namely, true positives, true negatives, false positives, and false negatives (FN). A TP represents the number of positive samples correctly categorized as positive, while a TN represents the number of negative samples correctly categorized as negative, a The number FP indicates how many negative samples have been mistakenly classified as positive, while the number FN indicates how many positive samples have been mistakenly classified as negative.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$



Accuracy: 92.72%

Fig. 14. Confusion Matrix

The accuracy of the proposed system was reported to be between 92-94%. This means that out of all the data points in the test set, the system correctly classified between 92-94% of them.

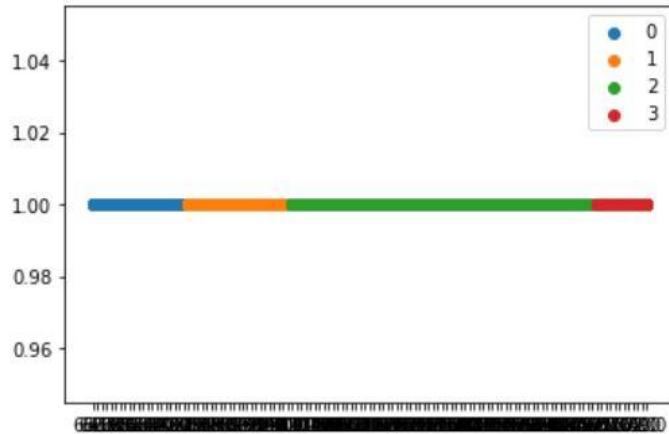


Fig. 15. Clusters

Clusters formed from activity dataset are mapped with the mood labels based on heart rates present in each cluster.

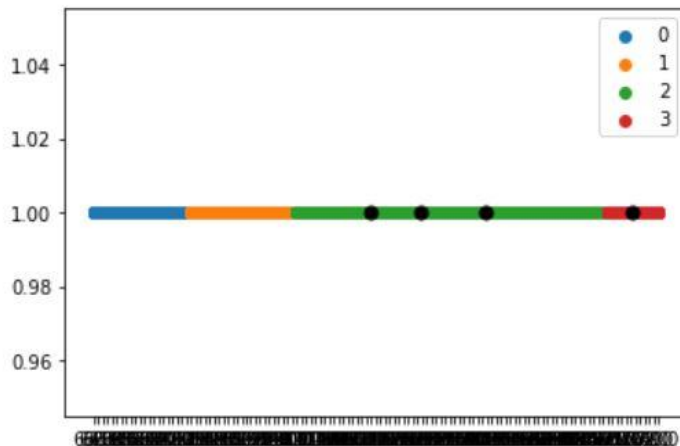


Fig. 16. Centroids of each cluster

## 5. Conclusion

A recommendation system using clustering and classification algorithms was developed. An average accuracy of 93% was achieved by using a stacking approach of ensemble learning. Clustering was performed on user activity data then clusters were mapped with appropriate mood. Final algorithm was developed based on mapped activity and classified songs data which recommends songs to the user based on his daily activities from time to time.

## 6. Future Scope

In the project, the user has to input his day to day schedule on that basis to get songs related to activity at a particular time. In the future the user's real time activity can be through wearable gadgets using heart-rate, blood pressure etc. Also this will increase the accuracy of

the clusters. The project currently predicts mood from current activity but in the future inter connections of current activity with previous activities and mood can be done to increase current mood accuracy and make songs more relatable. Also, in the future a feature of streaming songs can be added to make a better user experience.

## 7. References

- [1] Kim, Y., Aiello, L.M. & Quercia, D. PepMusic: motivational qualities of songs for daily activities. *EPJ Data Sci.* 9, 13 (2020). <https://doi.org/10.1140/epjds/s13688-020-0221-9>
- [2] Cristóbal Veas: Predicting the Music Mood of a Song with Deep Learning (2020). <https://towardsdatascience.com/predicting-the-music-mood-of-a-song-with-deep-learning-c3ac2b45229e>
- [3] Ankita Mahadik , Shambhavi Milgir , Janvi Patel , Vijaya Bharathi Jagan, Vaishali Kavathekar: Mood based Music Recommendation System ,IJERTV10IS060253 <https://www.ijert.org/mood-based-music-recommendation-system>
- [4] Jingzhou Yang: Personalized Song Recommendation System Based on Vocal Characteristics Volume 2022 | Article ID 3605728 | <https://doi.org/10.1155/2022/3605728>
- [5] Renata L. Rosa, Demóstenes Z. Rodríguez, and Graça Bressan: Music Recommendation System Based on User's Sentiments Extracted from Social Networks [https://www.researchgate.net/publication/283238023\\_Music\\_Recommendation\\_System\\_Based\\_on\\_User's\\_Sentiments\\_Extracted\\_from\\_Social\\_Networks](https://www.researchgate.net/publication/283238023_Music_Recommendation_System_Based_on_User's_Sentiments_Extracted_from_Social_Networks)
- [6] Speech Emotion Recognition Using Deep Learning Techniques: Ruhul Amin Khalil;Edward Jones;Mohammad Inayatullah Babar;Tariqullah Jan;Mohammad Haseeb Zafar;Thamer Alhussain. IEEE (2019) <https://ieeexplore.ieee.org/author/37594662300>
- [7] "Music Recommendation System Using Chatbot", International Journal of Emerging Technologies and Innovative Research (www.jetir.org | UGC and issn Approved), ISSN:2349-5162, Vol.9, Issue 5, page no. pph171-h178, May-2022, Available at : <http://www.jetir.org/papers/JETIR2205822.pdf>
- [8] Sricheta Parui; Abhiroop Kumar Roshan Bajiyya; Debasis Samanta; Nishant Chakravorty: Emotion Recognition from EEG Signal using XGBoost Algorithm <https://ieeexplore.ieee.org/abstract/document/9028978>
- [9] Shtwai Alsubai: Emotion Detection Using Deep Normalized Attention-Based Neural Network and Modified-Random Forest Sensors 2023, 23(1), 225; <https://doi.org/10.3390/s23010225>
- [10] Interacción '17: Proceedings of the XVIII International Conference on Human Computer Interaction Emotion detection: a technology review <https://dl.acm.org/doi/abs/10.1145/3123818.3123852>