International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 03 | Mar 2022 www.irjet.net p-ISSN: 2395-0072

Mehfil : Song Recommendation System Using Sentiment Detected

Sangita S. Keluskar [1], Vaishnavi L. Dhuri [2], Shreya S. Gonjari [3], Nidhi Sanghavi [4]

[1],[2],[3]Student, Department of Computer Engineering, Atharva College of Engineering, Mumbai
[4]Asst.Professor, Department of Computer Engineering, Atharva College of Engineering, Mumbai

---

Abstract - *Human Sentiments may be recognised through their facial expression, according to studies. As a result, the Mehfil project captures human expression and selects songs based on the emotions observed. Using one's system web camera, this can be done in real time. The sentiment is discovered using Machine Learning and distinguishing traits associated with the face. The Haar Cascade technique was used to recognise faces, which was then processed by a pre-trained lightweight model MobileNet V2. The study then displays a personalized playlist depending on the inferred sentiment. Furthermore, for the construction of the dataset and suggestion, this project makes use of the Spotify platform and API. The research project's main goal is to make the tedious and time-consuming operation of manually classifying music into separate lists easier by assisting in the creation of a relevant playlist based on an individual's emotional characteristics.*

Keywords: Spotify API, Flask, Haar Cascade, MobileNetV2, Sentiment Analysis, Client Id, Secret Id, Deep learning, playlist, authorization.

1.INTRODUCTION

Music is noted for its adaptability. It brings people of diverse ages, ethnicities, languages, tastes, political affiliations, and income levels together. Music is thought to be the universe's rhythm, capable of transporting one on any kind of long journey. Listening to the proper music can even help you feel better.

Melophiles frequently have a difficult time manually creating and organizing playlists, especially if they have a long list of favorites. It also gets tough for them to choose the appropriate music at the appropriate time to match their present mood. They now see the need for a mood detector as well as a recommender system to recommend tracks based on the mood. This reduces the amount of time and effort required to browse playlists.

In recent years, digital music technology has progressed, and with it, the development of music recommendation systems has progressed as well. Furthermore, music streaming services like Spotify, Gaana, Itunes, Amazon Songs, and Jio Saavn recommend music based on a user's recent listening history. For history-based suggestions, machine learning and deep learning algorithms are applied.

Content-based and collaborative-based recommendations are the two primary categories of recommendation systems. Users' current preferences are predicted by content-based algorithms based on their previous preferences. Collaborative-based systems estimate a user's preferences based on the preferences of other users who are similar to them. Most companies, such as Netflix and Hulu, adopt a hybrid method, in which recommendations are based on a combination of what a user enjoyed in the past as well as what other users like.

Recommendations based on collaborative filters generally follow two main kinds of approaches.

- User-based approach (or user-item):

  This approach looks for users with similar interests and behavior and considering what those similar users listened to, it makes a recommendation.

- Item-based approach (or item-item):

  This Approach, takes into consideration what songs the user has considered in the past and recommends new identical songs that the user can enjoy.

Facial expressions are more accurate than words in predicting human emotions. Happy, sad, angry, fearful, and neutral emotions can be used to offer music that are complementary to the user's mood. These feelings are difficult for a machine to grasp. A shift in emotion can be caused by a slight alteration in facial muscular contortions. Furthermore, because emotions are highly context-dependent, expressions of different or even the same individual for the same feeling may change. Even if we concentrate on the areas that give the most clues to emotions, such as the muscles surrounding the mouth and

eyes, extracting and categorizing these traits is a time-consuming process. With the help of Neural Networks and Machine Learning, these objectives may surely be met.

When it comes to recognising a song's mood, the music genre is insufficient because a single genre can encompass a wide spectrum of emotions. Audio features, on the other hand, play a vital role in music emotion identification. Happy, enthusiastic, and energetic emotions all have higher levels of intensity, timbre, pitch, and rhythm than quiet, contentment, and despair moods. To recommend songs to the user, the recommendation module integrates the y results of the emotion and music classification modules.

2. LITERATURE SURVEY

A great deal of previous research has focused on the field of face detection as it is the primary step in every vision-based automated system. These studies have been refined in terms of speed with the utilization of haar-features available in python along with the contribution of the Viola-Jones object detection framework. Outcomes of these frameworks, such as OpenCV, provide a variety of face classifiers [1]. Haar-Features are good at identifying edges and lines, which makes it especially functional in face detection.

In addition, a substantial amount of research on the significance of MobileNetV2 has been published. [2] Bharath Surendar I et al. presented "Facial Feature Analysis Using Deep Convolutional Neural Networks and HAAR Classifier for Real-Time Emotion Detection." The researchers behind this study developed a website that can predict a user's emotion. The goal of a Deep Convolutional Neural Network-based face emotion classifier is to use the deploy model in real-time.

Because there are so many different perspectives on how to classify songs based on one's mood, it necessitates a much more difficult grind. The extant literature suggests a self-contained and effective strategy for detecting the mood of any given piece of music. It also detects associated "emotions" by establishing a relationship between spectral and harmonic properties and human perception of music and moods. Rhythm, harmony, and spectral properties, to mention a few, are some of the features explored in the literature in order to perform song categorization [3].

There is a large volume of published studies describing the role of clustering algorithms. In this existing research, a comparison of two algorithms, K-Nearest Neighbor and ID3, has been done in the view of clustering songs. This research suggests an individual's mood by using K-Nearest Neighbor as the accuracy performance comparison and measurement of average classification time is better compared to ID3 on Thayer's model [4]. As feature selection for the clustering is necessary, the clustering is based on the music tracks' audio features involving acousticness, danceability, liveliness, energy, mode, tempo, danceability, speechiness, and valence [5]. This research provides an innovative approach for visualizing these clusters in order to represent the moods that result in the output of a customized playlist for an individual user.

The topic of music recommendation systems based on mood detection comprises a significant and expanding amount of literature. In a research [6] a web page is provided where the image is captured and transferred to the client-side and then by using a pre-trained VGG16 model, prediction was made and then the predicted output is passed to the client-side where songs were displayed using Spotify API.

Mood-based Music Recommendation System is a mobile application that uses a smartphone camera to capture a photo of the user and then generates a playlist depending on the mood identified [7]. Firebase was created where MP3 songs were uploaded manually in the storage section. After this, the Firebase database was linked to Android Studio. Finally, for the Android application, a UI was designed that grouped the file model methods and Firebase together.

Moodify is an emotion-based music player [7]. It's a music player that uses Chrome as a front-end to detect the emotions of the user. It also employs the Fisherface method to detect emotions, resulting in the generation of an XML model output. This model stores information about emotions. Front-end is implemented using JavaScript, which provides the manual selection of modes for the user. The Emo player is a music player where emotions are acknowledged based on a machine learning algorithm, Support Vector Machine (SVM) [8].

3. METHODOLOGY

Mehfil is a Song Recommendation system based on the sentiments detected of the user which primarily focuses on real-time sentiment analysis. It is composed of three main modules, firstly sentiment analysis, next is the music recommendation and finally the integration.

3.1 Sentiment Analysis Module

This module is built of two components:

- Face detection -

  Face detection is also known as facial detection, which is used to identify human faces in digital images. Additionally, it also helps to identify the part of an image that should be concentrated on. This project makes use of the Haar Cascade algorithm, which is an object detection algorithm used to distinguish faces in an image.

- Sentiment Analysis -

  To train the Sentiment Analysis model, the FER-2013 dataset from Kaggle is used [11]. From the dataset, a total of 25,102 images of the respective emotions were extracted. The images are in grayscale, each of size 48x48 pixels. For the classification of the images, a pretrained, lightweight model, MobileNetV2 from the Keras library, is used. MobileNetV2 is a CNN architectural model used for Transfer Learning. It uses depth-wise separable convolutions to build deep neural networks. This model is used as the base model, and the prepared dataset of five classes (happy, anger, neutral, sad, and fear) is fed into it.

3.2 Music Recommendation Module

This module is built of three components:

The Spotify API provided by one of the largest music streaming platforms, Spotify, is used to build the datasets of songs. It is an API based on simple REST principles. Metadata and other information about music artists, albums, and tracks in JSON format are returned by the Spotify Web API endpoints directly from the Spotify Data Catalog. Data resources are accessed via standard HTTPS requests in UTF-8 format to an API endpoint. In order to access Spotify API, a few steps needed to be followed. Firstly, registration of the app on the Spotify Developers profile is required. Followed by Authorization for which Client ID and Secret ID were mandatory. We have created two datasets viz. the labeled dataset which was used for training the model and the non-labelled dataset which was used to recommend songs.

- Non-Labeled Dataset:

  The Spotify API is used to produce a large dataset that is primarily targeted at the Indian market. Metadata, which includes artist, album, track name, and track id, was recovered from the extracted music data. The retrieved features include acoustic factors such as loudness, key, mode, and tempo, as well as psychoacoustic metrics like danceability, energy, instrumentals, liveness, and valence. The resulting dataset is saved as a CSV file. There are almost 200 thousand songs in the dataset, including regional classical Bollywood music. It is given as an input to the trained model for classification by mood. This dataset is used to recommend songs.

- Labeled Dataset:

  A labeled dataset was created which consists of all acoustic parameters like loudness, key, mode, tempo, and psychoacoustic parameters like danceability, energy, instrumentals, liveness, valence along with their metadata and mood label. A total of 13 different combinations of moods were obtained as shown in Fig -1. This dataset consists 2k+ songs. For creation of this dataset, we have made use of mood-based playlists which were already made by the Spotify users on Spotify.

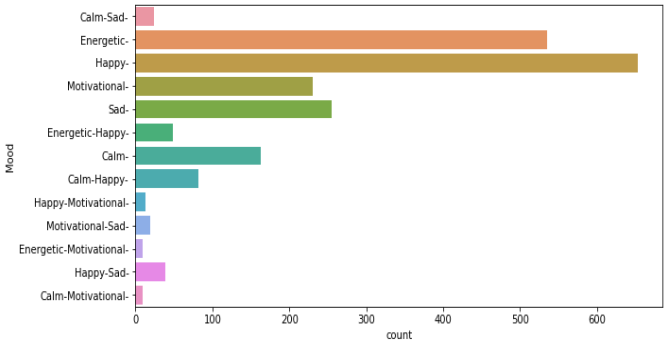

Fig -1: Distribution of Songs in Labeled Dataset

- Classification of songs:

  Multi label classification is used to overcome this issue because a single song can convey multiple emotions. The distinction is that any piece of music can be classified into one or more genres. As a result, each instance can be assigned to multiple categories, leading to multi-label classification challenges in situations where we have a number of target labels. As a result, each

song in the classified dataset was assigned one or more emotions.

Because the values of the dataset's columns, such as danceability, tempo, and so on, were highly variable, MinMaxScaler was employed to scale them. In addition, the distribution of songs in the 13 mood combinations obtained was severely skewed. As a result, the dataset was sampled using the Random Over Sampling technique.

A multiclass neural network was employed to classify songs in the non-labelled dataset. This was done using the KerasClassifier from the Keras Python library as an estimator. The features used for classification were danceability, energy, loudness, speechiness, acousticness, instrumentalness, liveness, valence and tempo. Initially, for testing the estimator, K-fold Cross Validation technique was used. Further, the labeled dataset was divided into training and testing sets. The model was trained using the training set. On evaluating the model using the testing set, the accuracy obtained was 59.12%. Finally, the prepared model was used to classify songs in the non-labeled dataset.

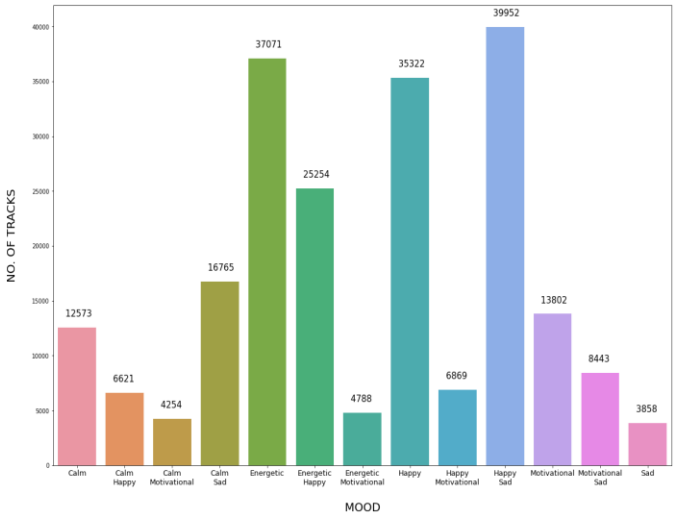

Fig -2: Distribution of Songs in Large Dataset

3.3 Integration

- Spotify Authorization:
  Dynamically accessing user history is important for which Spotify Authorization is performed. The process of allowing a user or application access permissions to Spotify data and features is known as authorization. One or more scopes determine who has access to the protected resources. Scopes allow your application to access certain functionality on behalf of a user (for example, reading a playlist, modifying your library, or just streaming). The access permissions that the user is requested to grant are determined by the scopes you establish during authorization. OAuth (Open Authorization) is a safe, industry-standard protocol that lets you authorize one application to connect with another on your behalf without disclosing your password. Instead of sending user credentials from app to app, OAuth uses access tokens (kind of like a special code) to pass authorization between apps over HTTPS. The process of allowing a user or application access permissions to Spotify data and features is known as authorization. The OAuth 2.0 authorization framework is used by Spotify. They give users of third-party apps peace of mind by ensuring that only the information they choose to share is shared.

- Connection:
  For the integration of the web application and python modules, this model has used Flask. On the client side, the image is captured in real time using the user's webcam. Furthermore, for sentiment analysis, the captured image will be further passed to the server-side, where at first, face detection will be done using the Haar Cascade algorithm, which returns the rectangular boundary coordinates of the detected faces. These detected face images will be transformed in grayscale and then converted to 224x224 size for feature point extraction. This is done since the MobileNetV2 model accepts a dataset of images of 224x224 size. Next, this image will be transferred to the saved model to predict the mood, and ultimately, the predicted mood will be transferred to the server along with the customized playlist of the songs. HTML CSS and JavaScript, user interface will be used which displays the image of the user along with their current mood and playlist.

4. HARDWARE AND SOFTWARE REQUIREMENTS

4.1 Hardware Requirements

One of the most common collections of requirements defined by any operating system or software application is

the physical computer resources, also known as hardware. The hardware requirements required for this project are:

- Development Operating System: Windows
- User System RAM: 2 GB and above
- Development RAM: 8 GB and above
- ROM: 10 GB HDD
- Dev Processor: Intel i5 and above
- Webcam (for testing on laptop/desktop)

4.2 Software Requirements

Software Requirements deal with defining software resource requirements and prerequisites that need to be installed on a computer to provide optimal functioning of an application. These requirements or pre-requisites are generally not included in the software installation package and need to be installed separately before the software is installed. The software requirements that are required for this project are:

- Anaconda, Python 3.8
- Python libraries
  - ➢ Open CV 4.0.1
  - ➢ Flask
  - ➢ Tensorflow 2.0
  - ➢ Keras
  - ➢ Spotipy
  - ➢ Scikit Learn Package
- Operating System: Windows (Vista/7 or above)
- Jupyter Notebook / Google Colab
- IDE (e.g. Visual Studio)

5. DESIGN DETAILS

5.1 Control Flow Diagram

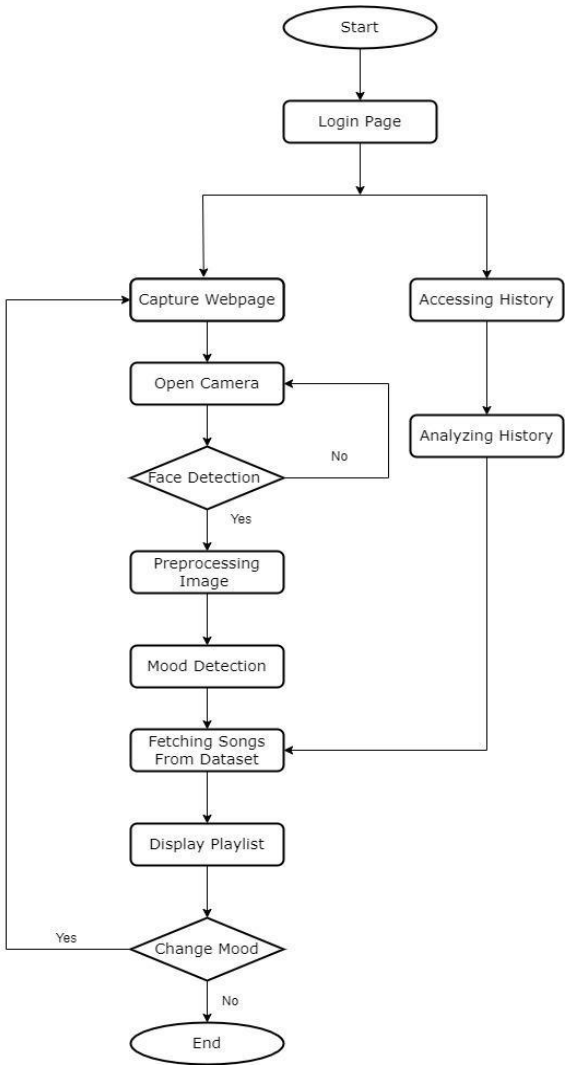

Fig -3: Control Flow Diagram of Mehfil

5.2 Design of System

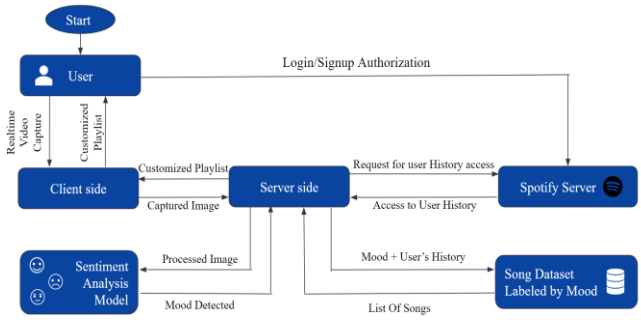

Fig -4: System Architecture of Song Recommendation System Using Sentiment Detected

6. RESULT ANALYSIS

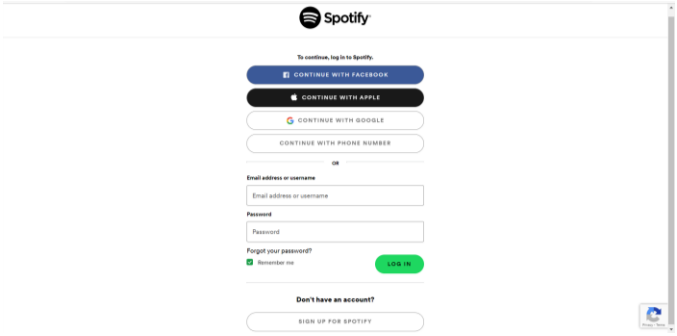

Fig -5: Spotify Authorization

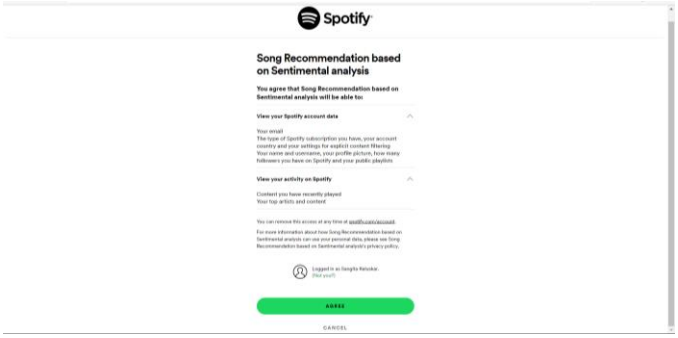

Fig -6: Spotify Authorization

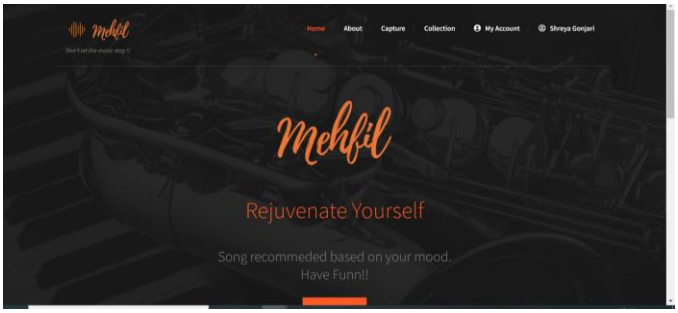

Fig -7: Home Page of Mehfil

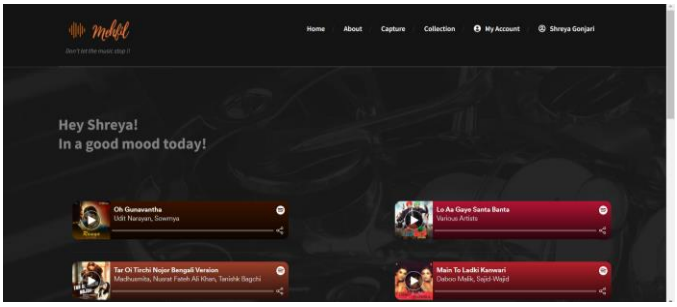

Fig -8: Output of Mood - Happy

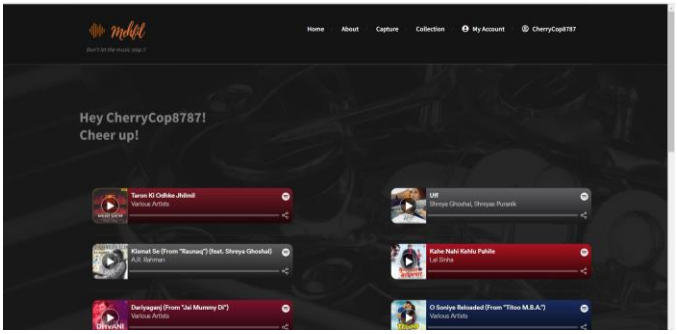

Fig -9: Output of Mood - Neutral

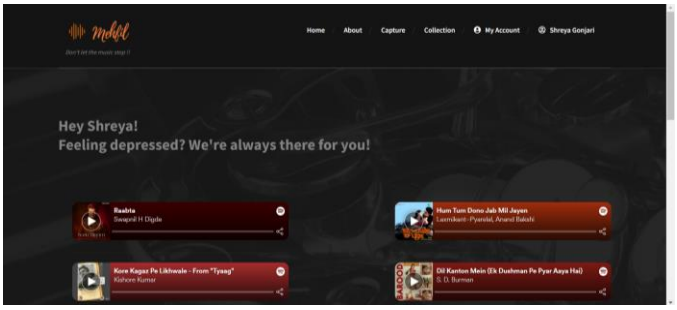

Fig -10: Output of Mood - Sad

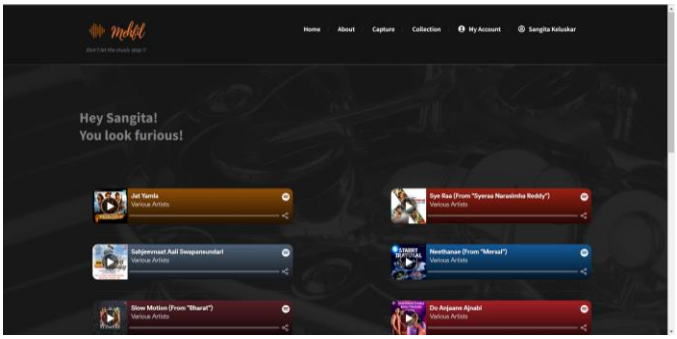

Fig -11: Output of Mood - Angry

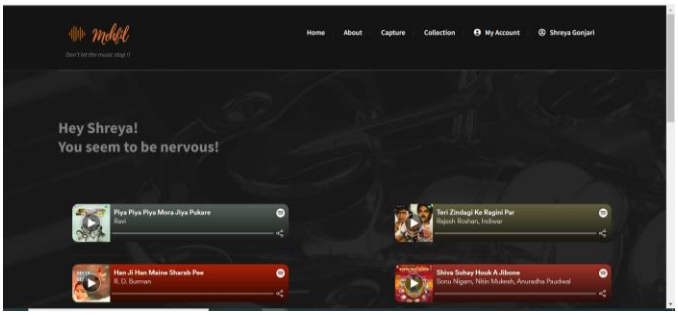

Fig -12: Output of Mood - Fear

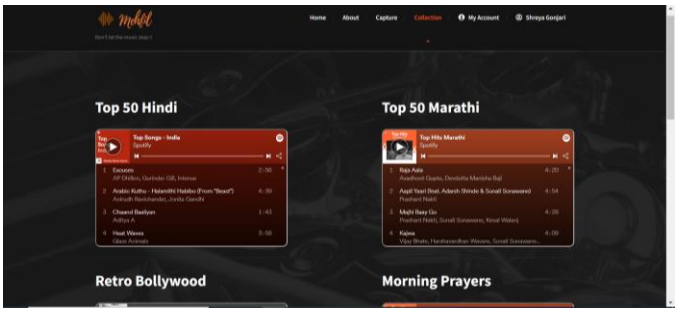

Fig -13: Collection Section

6. CONCLUSION

This research has been developed with an aim to contribute greatly in the field of machine learning and deep learning technology. Mehfil performs sorting out of the music based on one's emotions, such as whether they are happy or sad. Therefore, the main purpose of this research is to develop a web application in which a real-time image of the user can be captured and further based on the mood detected, a list of songs could be recommended. It is especially useful for reviving the user when he or she has free time and wants to listen to music that is appropriate for the scenario. Although detecting human emotions is quite a tedious task, machine learning makes it possible to a certain extent. This can be achieved by developing and training a machine learning model that can eventually detect a set of emotions based on one's facial expressions more precisely. The mood detection model, having an accuracy of 91%, is capable of detecting five moods accurately, viz. happy, sad, neutral, fear, and anger. In addition to that, it will have the ability to suggest a list of songs that would be suitable for the detected mood. It simplifies the work of the end-user by determining their emotion and suggesting a customized playlist based on the user's dynamic history accessed through Spotify API. This would assist melophiles in appreciating melodic tunes based on their recent played history.

ACKNOWLEDGEMENT

We owe sincere thanks to our college Atharva College Of Engineering, for giving us a platform to prepare a project on the topic "Mehfil: Song Recommendation System Using Sentiment Detected". We would further like to thank our Principal Dr. Shrikant Kallurkar for instigating within us the need for this research and giving us the opportunities and time to conduct and present research on the topic. We are sincerely grateful to Prof. Nidhi Sanghavi, our guide and Dr. Suvarna Pansambal, Head of Computer Engineering Department, for mentoring us throughout the duration of our project. Without their assistance, motivation, constant support and valuable suggestions, our project would not have developed the way it did.

REFERENCES

[1] Padilla, R., Filho, C. F. F. C., & Costa, G. F. (2012). Evaluation of Haar Cascade Classifiers Designed for Face Detection. International Journal of Computer, Electrical, Automation, Control and Information Engineering, 6(4). https://doi.org/10.5281/zenodo.1058133

[2] Surendar, B. I., Chakravarthy, P. S. S., Thangavel, R., & Swarnalatha, P. (2021). Facial Feature Analysis using Deep Convolutional Neural Networks and HAAR Classifier for Real-Time Emotion Detection, International Journal Of Engineering Research & Technology (IJERT), 10(5), https://www.ijert.org/research/facial-feature-analysis-using-deep-convolutional-neural-networks-and-haar-classifier-for-real-time-emotion-detection-IJERTV10IS050353.pdf

[3] Bhat, A., Amith, V., Prasad, N., & Mohan, M. (2014). An Efficient Classification Algorithm for Music Mood Detection in Western and Hindi Music Using Audio Feature Extraction. Fifth International Conference on Signal and Image Processing. http://dx.doi.org/10.1109/ICSIP.2014.63

[4] Sudarma, M., & Harsemadi, I. (2017). Design and Analysis System of KNN and ID3 Algorithm for Music Classification based on Mood Feature Extraction. International Journal of Electrical and Computer Engineering.7.486-495. http://dx.doi.org/10.11591/ijece.v7i1.pp486-495

[5] Bakhshizadeh, M., Moeini, A., Latifi, M., & Mahmoudi, M. T. (2019). Automated Mood Based Music Playlist Generation By Clustering The Audio Features. International Conference on Computer and Knowledge Engineering, 231-237, http://dx.doi.org/10.1109/ICCKE48569.2019.8965190

[6] Mahadik, A., Milgir, S., Jagan, V., Kavathekar, V., & Patel, J. (2021). Mood based Music Recommendation System. International Journal of Engineering Research & Technology, 10(6). https://www.ijert.org/research/mood-based-music-recommendation-system-IJERTV10IS060253.pdf

[7] Srinayani, M., Jahnavi, P. N., & Kavishree, S. (2021). Moodify: Smart Music Player based on Facial Emotions. International Research Journal of Engineering and Technology, 8(4). Retrieved from https://www.irjet.net/archives/V8/i4/IRJET-V8I479.pdf

[8] Hemanth, P., Adarsh, Aswani, C.B., Ajith, P., & Kumar, V. A., (2018). EMO PLAYER: Emotion Based Music Player, International Research Journal of Engineering and Technology, 5(4). Retrieved from https://www.irjet.net/archives/V5/i4/IRJET-V5I41065.pdf

[9] Nuzzolo, M. (2022). Music Mood Classification, Electrical and Computer Engineering Design Handbook. https://sites.tufts.edu/eeseniordesignhandbook/2015/mu

[10] Raut, N. (2018). Facial Emotion Recognition Using Machine Learning, Master's Projects, 632. https://doi.org/10.31979/etd.w5fs-s8wd

[11] Sambare, M. (2020). FER2013 [Dataset]. Kaggle. https://www.kaggle.com/msambare/fer2013

[12] Welcome to Spotipy! — spotipy 2.0 documentation (n.d). Spotipy.readthedocs.io. Retrieved from website: https://spotipy.readthedocs.io/en/2.19.0/