Music Recommendation System based on Emotion WebApp

Chandani Mourya , Prof. Pooja Singh
Student, Professor
Department of Computer Engineering
Terna Engineering College, Mumbai, India

Abstract : Songs have always been a popular choice to try and describe and understand human emotions. Emotion-based classification systems for data can be quite useful for people dealing with stress. In this paper, We introduce EmoMusic, a music player that plays and recommends songs based on the user's current mood. For identifying emotions, we used a convolutional neural network. By mapping the user's emotions to the mood type of song, our music player can identify human emotion through a webcam and suggest songs to the user. Our music player recommends the song as well as predicts the mood of the music based on its lyrics and audio.

***Keywords*— Convolutional Neural Networks; Deep Learning; Face Recognition, Song Recommendation**

I. INTRODUCTION

People frequently use their facial expressions to convey their feelings. A person's mood has long been known to be changed by music. It has been demonstrated by recent studies in the field of music psychology that listening to music causes listeners to experience a range of distinct emotions. The correlation between musical tastes and personality traits and moods has been found to be very strong. Brain regions that deal with emotions and mood control the metre, timber, rhythm, and pitch of music.

One can gradually quiet their thoughts and ultimately have a pleasing effect by capturing and recognising the emotion being shown by a person and playing music that suits that person's mood.

The oldest and most natural method of expressing emotions, moods, and sensations is through facial expressions. For the sake of this paper, we divide facial expressions into four main emotional groups: happy, sad, angry, neutral and surprise.

The goal of the project is to capture the emotion that a person expresses through their facial expressions and improve the user's mood by recommending and playing music that meets their needs as well as predict the mood according to the lyrics and audio of song.

II. LITERATURE SURVEY

Parul Tambe et al. [1] proposed an approach that automated user-music player interactions, which learnt all of the user's preferences, moods, and activities and provided song selection as a result. The device captured users' varied facial expressions in order to analyse their emotions and anticipate the musical style based on those emotions.

S. Deepika, K. A. Indira, and Dr. Jesline[2] employed Convolutional neural networks to recognise emotions and play music in response. The music fan can find it helpful to separate the tunes and play them according to their mood. CNN has built a model that plays a song in response to the user's emotions, condensing many actions including collecting, detecting, and classifying the emotion into a single step.

Jayshree Jha et al. [3] presented an emotional music player using image processing. This demonstrated how numerous algorithms and strategies that were proposed by many writers in their research may be employed for linking the music player together with human emotions. It has therefore assisted in minimising the user's efforts in developing and managing playlists and giving exceptional service to music listeners by bringing them the most appropriate song in accordance with the user's present expression.

Prof. Nutan Deshmukh et al. [4] centred on developing a system that automatically detects a user's emotion using a camera and a camera-based technology, and then automates the outcome. This algorithm records the user's mood after each predetermined interval of time because the user's mood might not remain the same after some time and might even change. The proposed technique, which was superior to previously existing algorithms and lowers the cost of creating, on average computed estimation takes about 0.95 to 1.05 seconds to build an emotion-based music system.

III.METHODOLOGY

1.Music Recommendation

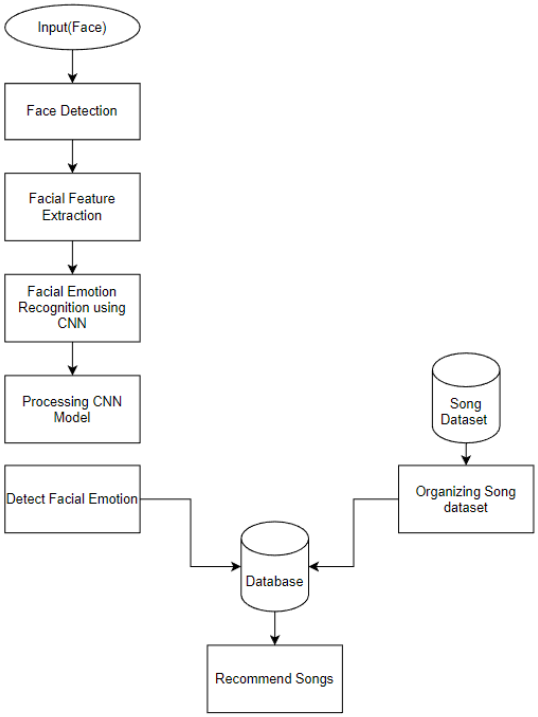

Fig. 1 Facial Emotion based Music recommendations Flowchart

1.1 Emotion Module

1.1.1 Dataset Description

The dataset we used for training the model is from a Kaggle Facial Expression Recognition Challenge, FER2013. Grayscale portraits of faces measuring 48 x 48 pixels make up the data. Each of the seven emotion categories—angry, disgust, fear, happy, sad, surprise, and neutral—is represented by one of the faces. Four different emotions—angry, happy, sad, neutral and surprise—were used in this study. There are 26,217 photos in all that go with these feelings. The photographs were divided into four categories: happy (8,989 samples), sad (6,077 samples), neutral (6,198 samples), angry (4,953 samples) and surprise(4002).

1.1.2 Model Description

The features of the user image are evaluated using a multi-layered convolutional neural network [6], [7]. An input layer, a few convolutional layers, ReLU layers, pooling layers, several dense layers (also known as fully-connected layers), and an output layer make up the convolutional neural network. These layers are sequentially and linearly stacked. The model achieved an accuracy of 74.2%.

1.2 Recommendation Module

1.2.1 Dataset Description

The dataset includes 390 songs divided into five moods. Class 0 has 100 songs, Class 1 has 93 songs, Class 2 has 100 songs, Class 3 has 95 songs, and Class 4 has 97 songs. This is how the songs are distributed. The songs' labels were done by hand.

Songs of anger make up Class 0, glad and joyous songs make up Class 1, neutral songs make up Class 2, sorrowful songs make up Class 3, and startled songs make up Class 4.

1.2.2 Model Description

To assess the song's mood, a multi-layered neural network was trained. A thick output layer, numerous hidden layers, and an input layer are all present in the network. The input layer's dimensions are set and predetermined. The ReLU procedure is used to add nonlinearity to the dataset. This guaranteed that the model also functions well in real-world circumstances. We are able to combine features in the hidden layer, which is a conventional multi-layer perceptron, improving the classification accuracy. A softmax activation function was utilised in the output layer to create the output as a probability for each mood class.

The result obtained from above was very promising. The Music player takes Image as input from either Webcam, URL or uploading in any image format and perform the emotion detection on the image to recognise the emotion of user and according to the emotion a playlist of relevant songs is generated. Similar songs are grouped together while generating the playlist of song to be recommended.

The accuracy is 74.5%

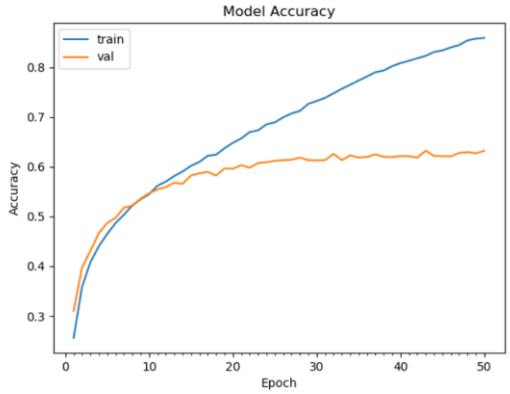

Fig. 2 Accuracy

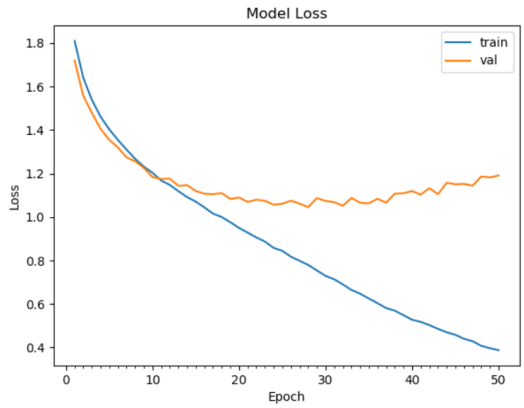

Fig. 3. Loss

2. Lyrics Mood Prediction

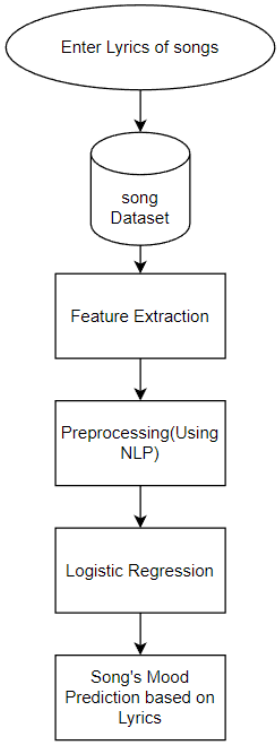

Fig. 4. Music's mood prediction based on Lyrics Flowchart

2.1 Dataset Description

The dataset used to train the model for the prediction of lyrics'mood is from githubusercontent website named train_lyrics_1000.csv and from github website valid_lyrics_200.csv

2.2 Model Description

To determine a song's mood from its lyrics, With the use of logistic regression, the model was trained. We processed the lyrics data using NLP to perform stemming, remove stop words, and other tasks. Data has been

vectorized, and the vectorized data has been saved in pkl format. We compared Logistic regression to Naive Bayes and Random Forest classifier and found that it produces the best results. The classification model was trained and saved in pkl format. We performed lyrics prediction using both models.

3. Audio Mood Prediction

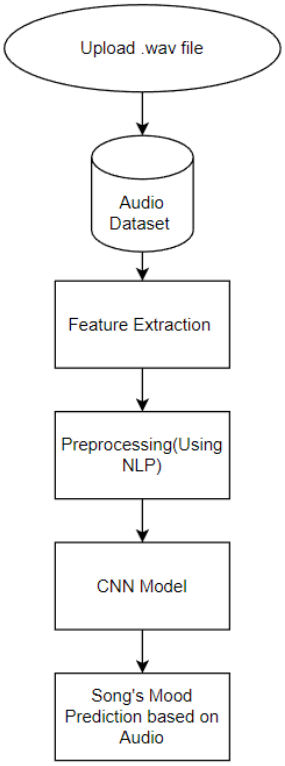

Fig. 5. Music's mood prediction based on Audio Flowchart

3.1 Dataset Description

We have used 5252 samples from the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) dataset the Toronto emotional speech set (TESS) dataset The samples include:

1012 song files and 1440 speech files are provided by RAVDESS. 24 professional actors, 12 male, and 12 female are recorded in this dataset vocalizing two lexically similar statements with neutral North American accents. Both speech and music can display a range of emotions, including happiness, sadness, anger, fear, surprise, and disgust. Ten times each, the emotional validity, intensity, and sincerity of each file were appraised. 247 people representing the untrained adult research participants from North America gave their ratings. 72 additional participants supplied the test-retest information. High levels of test-retest interrater reliability, interrater reliability, and emotional validity were reported. Our study and the validation data are available for download from PLoS ONE.

2800 TESS files. Two actresses (26 and 64 years old) recited a set of 200 target words in the carrier phrase "Say the word _____," and recordings of the set illustrating each of the seven emotions were made. (anger, disgust, fear, happiness, pleasant surprise, sadness, and neutral). There are 2800 total stimuli. From the Toronto region, two actresses were sought after. Both actresses have studied at universities, had musical training, and English is their first language. According to audiometric testing, the thresholds of both actresses fall within the normal range.

3.2 Model Description

We compared CNN approach with accuracy of 85.4% and decision tree classifier with accuracy of 68% in order to determine the mood of a song based solely on its audio, and found that CNN approach provides greater accuracy. We used the CNN technique to train the model, and we stored it in h5 format. The model is employed to forecast the song's mood based on the audio file.

Result:

Table 3.1: performance of features

Features	Accuracy
Image Detection and song Recommendation	0.74
Lyrics detection	0.68
Audio Detection	0.85

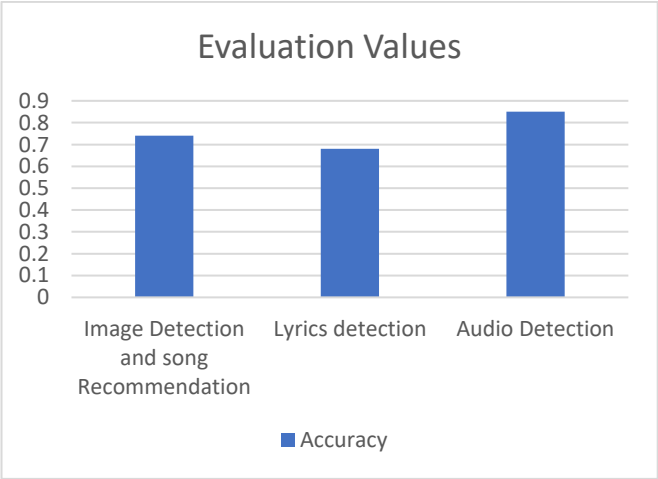

Fig. 6. Graphical presentation of performance of Features

This graph shows what percentage of accuracy are present in our models. The System work quite well with the dataset that we have chosen to train our all three models. We have mainly considered the Facial features to detect Emotions, song lyrics and tempos and music of songs for the developing of all three features respectively.

So overall our system performance accuracy is 75%.

IV. CONCLUSION

A detailed analysis of the literature reveals that there are various ways to put the Music Recommender System into practise. In this system, we give a general explanation of how music can effect a user's mood as well as advice on how to pick the best music tracks to lift a user's spirits. The technology in place is capable of identifying the user's emotions. The system was able to identify happy, sad, angry, neutral, or shocked emotions. Our system is also able to identify the mood of the song on the basis of lyrics as well as audio. The suggested approach presented the user with a playlist of music matches that corresponded to the user's emotion after identifying it.

We understand that there is room for development.. It would be intriguing to examine how the algorithm performs when all seven fundamental emotions are taken into account. Additional songs from various languages and geographical locations might also be added to strengthen the suggestion system. In a later piece of work, we intend to solve these problems.

REFERENCES

[1] ParulTambe, YashBagadia, Taher Khalil and Noor UlAin Shaikh 2015 Advanced Music Player with Integrated Face Recognition Mechanism International Journal of Advanced Research in Computer Science and Software Engineering.

[2] Vivek JD, Gokilavani A, Kavitha S, Lakshmanan S and Karthik S 2017 A novel emotion.

[3] S. Deepika, K.A. Indira, Dr. Jesline, "A machine learning Based Music player by Detecting emotions" recognition based mind and soul-relaxing system In2017 International Conference on Innovations in Information, Embedded and Communication Systems 1-5 IEEE.

[4] Rabashette MD, Tale MR, Hinge MA, Padale MK, Chavan MR and Deshmukh N Emotion

[5] Ram kumar madupu, Chiranjivi kothapalli vasanthi, yarra sonti harika,"Automatic Human Emotion Recognition system using Facial expressions with convolutional Neural Network

[6] Konstantious pyrovolkis, paraskevi Tzouvali, George stamou, "Mood detection analysing lyrics and audio signals bassed on deep learning architectures"

[7] Miao jiang, ziyi yang, chan Zhao, "What to play next? A RNN-based music recommendation system",

[8] Hui zhang, Kejun Zhang, Nick Bryan-Kinns, "Exploring the emotional preference of music for music recommendation in daily activities"

[9] Rahul Chauhan, Kamal Kumar Ghanshala, R.C Joshi, Convolutional Neural Network (CNN) for Image Detection and Recognition"

[10] G Jawaherlalnehru , S Jothilakshmi, "Music Emotion Classification using Logistic Regression and Random Forest".