**ACKNOWLEDGEMENT**

With reverential pranams, we express our sincere gratitude and salutations to His Holiness **Dr. Sree Sree Sivakumara Swamigalu**, Founder President of Sree Siddaganga Education Society and **Sree Sree Siddalinga Swamigalu**, President of Sree Siddaganga Education Society for their unlimited blessings.

First and foremost, we wish to express our deep sincere feelings of gratitude to our institution, **Siddaganga Institute of Technology**, for providing us an opportunity for completing our project successfully.

We are grateful to **Dr. M N Channabasappa**, Director, and **Dr. Shivakumaraiah**, CEO, Siddaganga Institute of Technology, Tumakuru for their kind cooperation and encouragement.

We express our kind thanks to **Dr. Dinesh S V,** Principal, Siddaganga Institute of Technology, Tumakuru for his encouragement and support to complete our mini project.

We express our heartfelt thanks to **Dr. A S Poornima**, Professor and Head, Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumakuru for her suggestions and advice.

We express our gratitude and humble thanks to our project guide, **Dr Shobha K**, Assistant Professor, Department of Computer Science and Engineering, Siddaganga Institute of Technology, Tumakuru for guiding and facilitating to complete our Mini-project work successfully.

We are conscious of the fact that we have received cooperation in many ways from the **Teaching, Technical and Supporting Staffs** of the Department of Information Science and Engineering and we are grateful to all for their cooperation.

We express heartfelt gratitude to our **Parents and Friends** for their constant moral support and encouragement throughout this work.

## ABSTRACT

Songs, as a medium, have always been a popular choice to depict human emotions. We validate our models by creating a real-time vision system that accomplishes the tasks of face detection and emotion classification simultaneously in one blended step using our proposed CNN architecture. Reliable emotion-based classification systems can go a long way in facilitating emotions. However, research in the field of emotion-based music classification has not yielded optimal results. In this work, we present an effective cross-platform music player, Emotion Music Player, which recommends music based on the real-time mood of the user. Emotion Music Player provides smart mood-based music recommendations by incorporating the capabilities of emotion context reasoning within us. The Emotion Module takes an image of the user as input and makes use of machine learning algorithms to identify the mood of the user.

In this work, computer vision components are used to determine the user's emotion through facial expressions. Once the emotion is recognized, the system suggests a play-list for that emotion, saving a lot of time for a user over selecting and playing songs manually.

The objective of this project is to detect emotion and recommend songs or music to be played based on the detected emotion.

## Contents

1. Introduction 6
2. Literature Survey 7
3. Requirement Specification 9
4. Design and implementation 11
5. Results 17
6. Conclusion 19
7. References 20

## Introduction

Music is an important entertainment medium. With advancement of technology, the optimization of manual work has gained a lot of attention. Currently, there are many traditional music players that require songs to be manually selected and organized. User, have to create and update play- list for each mood, which is time consuming. Some of the music players have advanced features like providing lyrics and recommending similar songs based on the singer or genre.

Although some of these features are enjoyable for user, there is room to improve in the field of automation when it comes to music players. Selecting songs automatically and organizing these based on the user’s mood gives user’s a better experience. This can be accomplished through the system reacting to the user’s emotion, saving time that would have been spent entering information manually. Emotions can be expressed through gestures, speech, facial expressions, etc. For the system to understand a user’s mood, we use facial expression. Using the device’s camera, we can capture the user’s facial expression. There are many emotion recognition systems which take captured image as input and determine the emotion.

Recent [studies](https://www.tamucc.edu/science/~cams/projects/523.pdf) confirm that humans respond and react to music and that music has a high impact on person’s brain activity. The average American listens up to four hours of music every day. People tend to listen to music based on their mood and interests. This project focuses on creating an application to suggest songs for user based on their mood by capturing facial expressions.

Smart gadgets people carry everyday with them can obtain lot of data. Data from fitness band, emotion detection from face captured by smartphone, activities detected by the sensors, further can be used for various applications. One of the very precise and straight approach to detect mood is using human facial expressions. Most of the time, emotion is revealed by face itself. Industries have their trained model for the emotion extraction.

Machine learning can be used for classifying music into set of particular emotions. Once, all this data is present, user can be studied about his/her preferences and habits of listening by time, mood, activity, etc. Training on this data can generate better playlist for future listening. Objective behind this work is to let daily factors get considered for better music recommendation

Everyone wants to listen music of their individual taste, mostly based on their mood. Average person spends more time to listen music. Music has high impact on person brain activity. User always face the task to manually browse the music and to create a playlist based on the current mood. This project is very efficient which generate a music playlist based on the current mood of user. However, the proposed existing algorithms in use are comparably slow, less accurate and sometimes even require use of additional hardware like EEG or sensors. Facial expression is a easy way and most ancient way of expressing emotion, feelings and ongoing mood of the person. This model based on real time extraction of facial expression and identify the mood.

## Literature Survey

The Following Research Papers were referred for getting ideas about the project, features and implementation

The external body part is a vital part of somebody's body and it particularly plays a vital role in knowing somebody's mood. Extracting the desired input from the external body part will currently be done directly employing a camera.[1] This input will then be employed in some ways. One in all the applications of this input is for extracting the knowledge to deduce the mood of a person. This information will then be wont to get an inventory of songs that befits the mood derived from the input provided earlier.

This eliminates the long Associate in the nursing tedious task of manually segregating or grouping songs into totally different lists and helps in generating an acceptable listing that supported somebody's emotional options. Varied algorithms are developed and planned for automating the listing generation method. Countenance-based mostly Music Player aims at scanning and deciphering the information and consequently, making a listing based on the parameters provided. The scanning and deciphering include audio feature extraction and classification to induce an inventory of a song's happiness to an analogous genre or to induce an inventory of comparable sounding songs.

The easiest way to express emotions for humans is using facial expressions. We humans, often use nonverbal cues such as hand gestures, facial expressions, and tone of the voice to express feelings. The proposed approaches have focused only on the some of the basic emotions. For this reason, facial features have been categorized into two major categories such as Appearance-based feature extraction and Geometric based feature extraction by Zheng et. al [1]. The paper by Hafeez Kabili et al [2] suggested the problem of the existing methods to handle only deliberately displayed and exaggerated expressions of prototypical emotions despite the fact that deliberate behaviour differs in visual appearance, audio profile, and timing from spontaneously occurring behaviour, by taking efforts to develop algorithms that can process naturally occurring human affective behaviour have recently emerged. They also introduced and researched this recent information and discussed human emotion perception from a psychological perspective. Nikhil Zaw are et al [3] stated that it is very time consuming and lengthy task to create and manage large playlists and to select songs from such playlists. Therefore, it would be of great use if the music player itself selects a song according to the current mood of the user using an application to minimize the efforts of managing playlists. In their paper, they stated a way to detect the mood of the user automatically and generate playlist of songs which is suitable for the user’s current mood. The photo is captured using webcam and that photo is passed under various steps to detect the mood or emotion of the user.

3**.** Renuka R Landhi et al. [1] proposed a paper which focused on the study of changes in the curvatures of the face and the intensities of the corresponding pixels. The author used Artificial Neural Networks (ANN), which was used to classify the emotions. The author also proposed various approaches for a playlist. Zheng et al. [2] proposed two significant categories for facial feature extraction, which included Appearance-based feature extraction and geometric based feature extraction, which included extraction of some essential points of the face such as mouth, eyes, and eyebrows. Nikhil et al. [3]. determines the mindset of the user by using facial expression Humans often express their feeling by their expressions, hand gestures, and by raising the voice of tone but mostly humans express their feelings by their face. Emotion-based music player reduces the time complexity of the user. Generally, people have a large number of songs on their playlist. Playing songs randomly does not satisfy the mood of the user. This system helps user to play songs automatically according to their mood. The image of the user is captured by the web camera, and the images are saved. The images are first converted from RGB to binary format. This process of representing the data is called a feature-point detection method. This process can also be done by using Haar Cascade technology provided by Open CV. The music player is developed by using a java program. It manages the database and plays the song according to the mood of the users. Zeng et al. [5] researched various advances in human affect recognition. He focused on various approaches that can handle audio and/or visual recordings of displays of affective states. The paper provides a detailed review of audio/visual computing methods. The effect is described as a prototype of emotion categories which include happiness, sadness, fear, anger, disgust, and surprise. This paper focused on discussing the challenges in computing methods for the development of automatic, spontaneous affect recognizer, which helped in emotion detection. It also identified some problems that have been missed or avoided in unit-modal posed emotion recognition.

4. Currently there are different methodology proposed by researchers to classify the emotional state of human behaviour. We have only focused on some of the basic emotion of human. A precise and efficient approach for examine the extracted facial expression was developed by Renuka R. Londhe et al. These documents mainly focused on the study of changes in the facial curve and it also focus on the intensity of the corresponding pixels. The artificial neural networks (ANN) were used to classify the characteristics extracted in 6 main universal emotions such as anger, disgust, fear, happiness, sadness and surprise. A scaled conjugate gradient back propagation algorithm correlated with a two-layer neural network was used and achieved a detection rate of 92.2%. In order to reduce the human effort and time required to manually separate songs from a playlist, different approaches have been proposed in correlation with different classes of emotions and moods.

# Thayer [16] proposed a very useful two-dimensional model (stress energy v / s), plotted two axes and whose emotions are represented by a two-dimensional coordinate system based on two axes or on the four quadrants, which is represented by the two-dimensional diagram be formed. , The musical mood names and AV values of a total of 20 subjects were tested and analyzed in the work of Jung Hyun Kim [7]. Based on the results of the analysis, the aircraft AV was divided into 8 regions (clusters), which illustrate the mood using an efficient data mining algorithm for k-means clusters.

**Requirement Specification**

# Functional

1. Our project provides a platform for easy interaction as easy as plug and play.

2. Provides separate modes for playing music.

3. Users can manually play songs just by clicking on the songs

## Software Requirements:

The software requirements of the project include:

* Front End: HTML, CSS, JS, Bootstrap
* Python
* Python Libraries: Pandas, NumPy, Scikit-Learn.

## Front End:

* 1. **HTML: -** The **HyperText Markup Language** is the standard markup language for documents designed to be displayed in a web browser. Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. We have used 5 webpages for our website.
  2. **CSS: - Cascading Style Sheets** (**CSS**) is a style sheet language used for describing the presentation of a document written in a markup language such as HTML.
  3. **JavaScript (JS): - JS** is a programming language that conforms to the ECMAScript specification. As a multi-paradigm language, JavaScript supports event-driven, functional, and imperative programming styles.
  4. **Bootstrap: - Bootstrap** is a free and open-source CSS framework directed at responsive, mobile-first front-end web development.

**Python: - Python** is an interpreted high-level general-purpose easy-readable programming language. Its language constructs and its object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. We have used python to compute calculations for the input data with the logical approach and rendered the result back.

## Libraries: -

* 1. **Pandas: pandas** is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.
  2. **NumPy: NumPy** is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

## Hardware Requirements:

The minimum hardware Requirements of the project is a laptop with:

* Processor: i3
* RAM: 4GB
* LAN/Wi-Fi Support
* Hard disk :256GB

**Non-Functional**

**Availability:** Our application is made available all the time and the users can access the application through mobile or pc.

**Reliability:** Once the user opens the application it is guaranteed that the song will be played

in the respective mode.

## Design and Implementation

**Detecting emotions**

Collecting data involves data set need to train the model

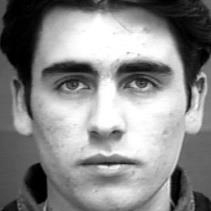
1. Angry

2.Sad



3. Happy

1. Neutral



**Facial expression detection in the Fisher's face works with the help of trained models. The reason behind this is to allow the user to take the dataset according to their use. Suppose if we take a huge amount of dataset of around 25-30k it will give nice accuracy no doubt but if the situation is like that the user of the devices, are a few people. Now in such conditions, if we take some precise dataset with around 400-450 images as input related to the user then it will also give good accuracy with the benefit of less amount of dataset and less storage on memory to operate. As well as a small memory of data gives output fast which results in quick response time. Here we first tried with the Cohn-Kaneda dataset then we made some classification in the as our need make it to train our model.**

**Implementation:**

Loading and saving trained model

For training, we have used Fisherface method of cv2 library.



For training data model, we have made a python code which grab all the classified images from folders and map it with its emotion. These data we at an instance stored in dictionary and then use. train method to train model.



To save the model for later use we have implemented. save method.



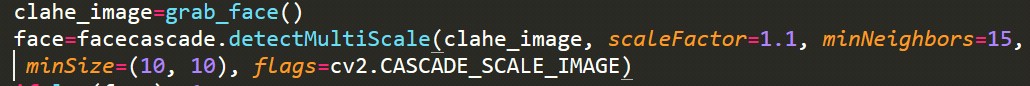
Now at the detection time first we have load model in memory using. read method.

Prediction of result is based on the prediction and confidence value which. predict method return.



Haarcascade model

Haarcascade model is precise face detection trained model which is provided by Open-cv. It returns the co-ordinates in terms of (x, y) at (left, bottom) of face frame and its width and height from those co-ordinates.



As here in the. detectMultiScale() method it is capable of detect multiple faces and it return an array of all the faces(co-ordinates) as an element.

The arguments have set according to the threshold what we need for our checking purpose. We have set it such like it doesn’t affect our model accuracy.

**Result Calculation**

In our model we have not stick on one image for testing, While the code will run it will take around 10 images in a short time (1-2 sec) and for all those images it will compute result and according to the average value of that it will give result. Apart from that we have make two codes one work on single face at a time while another work with multiple faces in the image.

**Machine Learning**

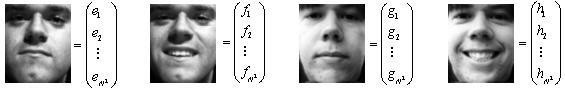
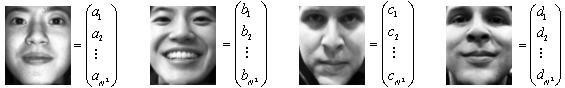
Fisherface ML algorithm

Fisherface algorithm is an algorithm which work on the basis of LDA and PCA concepts. Linear discriminant analysis (**LDA**) is a supervised Learning method of machine learning. Now supervised Learning is that where we use such data whose answer is also given to the model to learn it. It works on the concept of dimensionality reduction. Which reduce the execution time among classification?

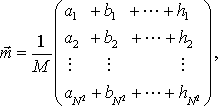
**Principal Component Analysis** (**PCA**) is a one kind of conversion from correlated variables to uncorrelated in the form of mathematical values.

It is mostly used for the observing data and from that by some probabilistic calculation generate models. The flow of Fisherface is like it takes classified images then it will reduce the dimension of the data and by calculating its statistical value according the given categories it stores numeric values in .xml file. While prediction it also calculates the same for given image and compare the value with the computed dataset values and give according result with confidence value.

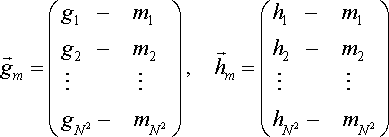
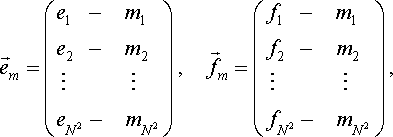
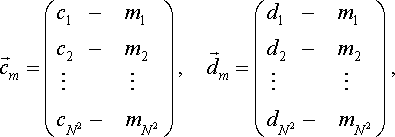
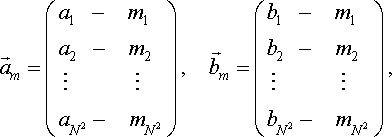
PCA Algorithm

Conversion training image **1, 2, . .. 𝑚 with size N x N into vector form with length size N ^ 2

Calculate the average of all face images



Calculate Matrix A with the formula



*A*   *a*  *m***,***b*  *m***,****,** *h*  *m* **[***am* **,** *bm* **,** **,** *hm* **]**

**Resizing images**

Whatever the image we have chosen for dataset it mostly related to the size which can give a precise output. The size is chosen such like the model can able to easily distinguish face from image by haarcascade model. And the size what we get from real time scan is not always same as data (very less difference) so, we resize it to the exact model data size. In our case we have chosen 350\*350.

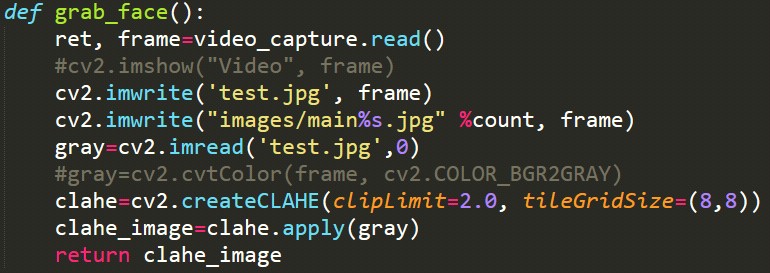


Here In this method, we have implemented the cropping of image by given parameters of haarcascade by clahe\_image[] and use of cv2’s method .resize() to the given size. Finally, we have stored those images in dictionary and after some count (=10) take it to check result.

Gray scaling images

It was the need for the method and because of it’s contrast and shaded face, it result in benefit for algorithm to get output.

Face detection



As the given in the code grab\_face() methods uses to get the images and do all operation and finally return cropped ,grayed face value in dictionary.

Train and predict methods

This code is use to get prediction and confidence value for given amount of image. Then get the max function with obtained output and final result is shown to the user.

**Playing music**

Detected emotions

We have implemented the linking of python with JavaScript through eel library. Which provide us the privilege to access python methods from JavaScript as well as vice versa. Here the striating flow will be in python code as the library is implemented in python then it transfers the control to html, JS. And according to the result we show emoticons.



Sad happy angry neutral

According to which we can classify emotion directory for playing song we have chosen these 4 emotions.

Methods for playing songs

In JavaScript file we have implemented too much methods for the switching of song.

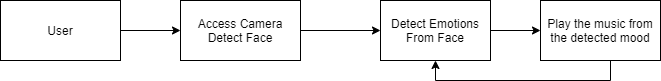
* 1. Queue
  2. Based on Emotion
  3. Random

In the first one as queue works it has been implemented. In second one we call python code to get emotion from user’s facial expression and according to that chosen next song which is also randomly and played it. In third one we directly used random function and all the methods are dynamic it can handle as change in number of songs accordingly.

HTML, CSS and JS concepts for online music player.

As we know the css give a great look to communicate and through JS we can interact with user and not look like complicated program run at console and it also give user privilege to choose any song to play.

## System Diagram:



**Fig (1) Flow Diagram**

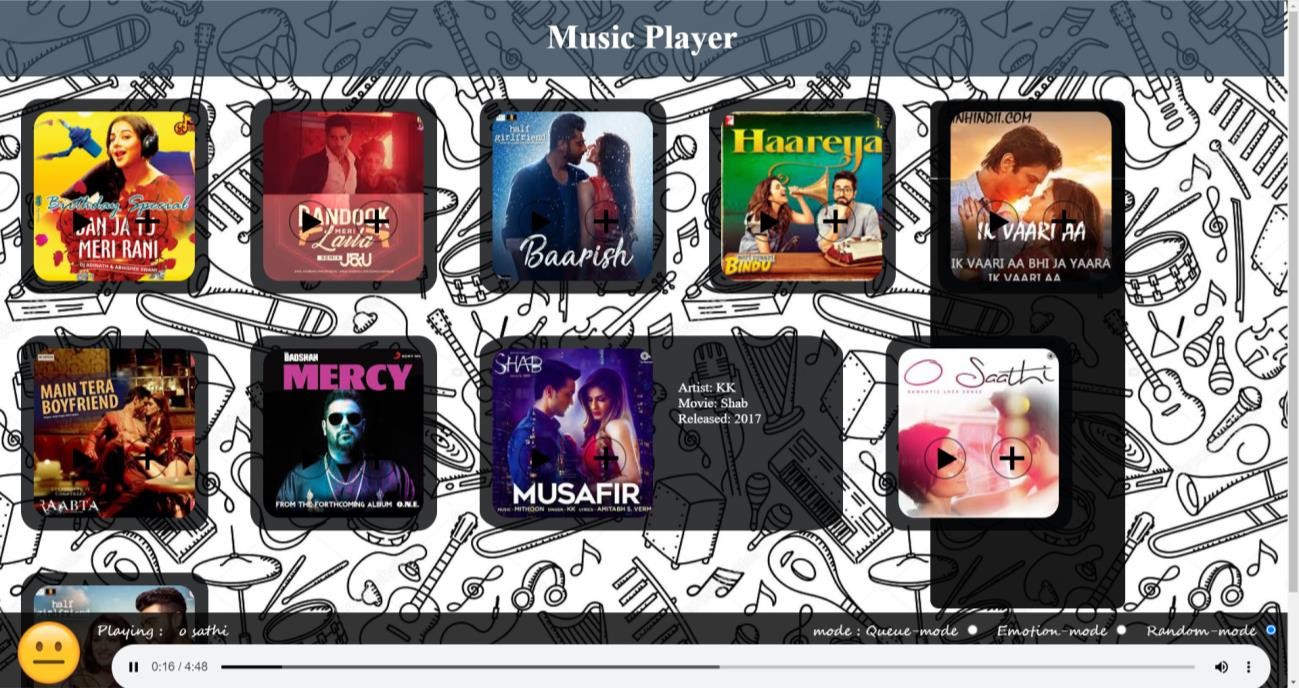
User allows to access the camera and once the permission is granted, the emotion-based music

player detects the face and emotions are computed.

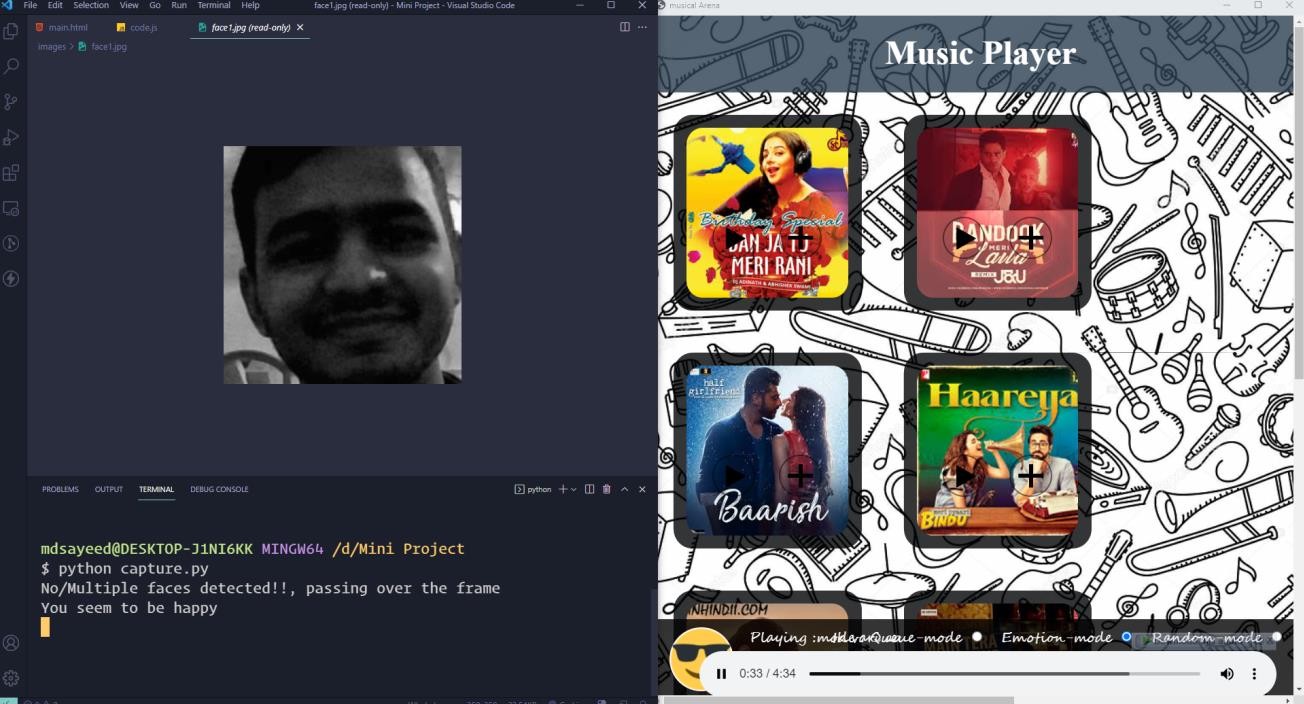
Based on these emotions, our model analyzes and recommends the songs from the list.

**Results**

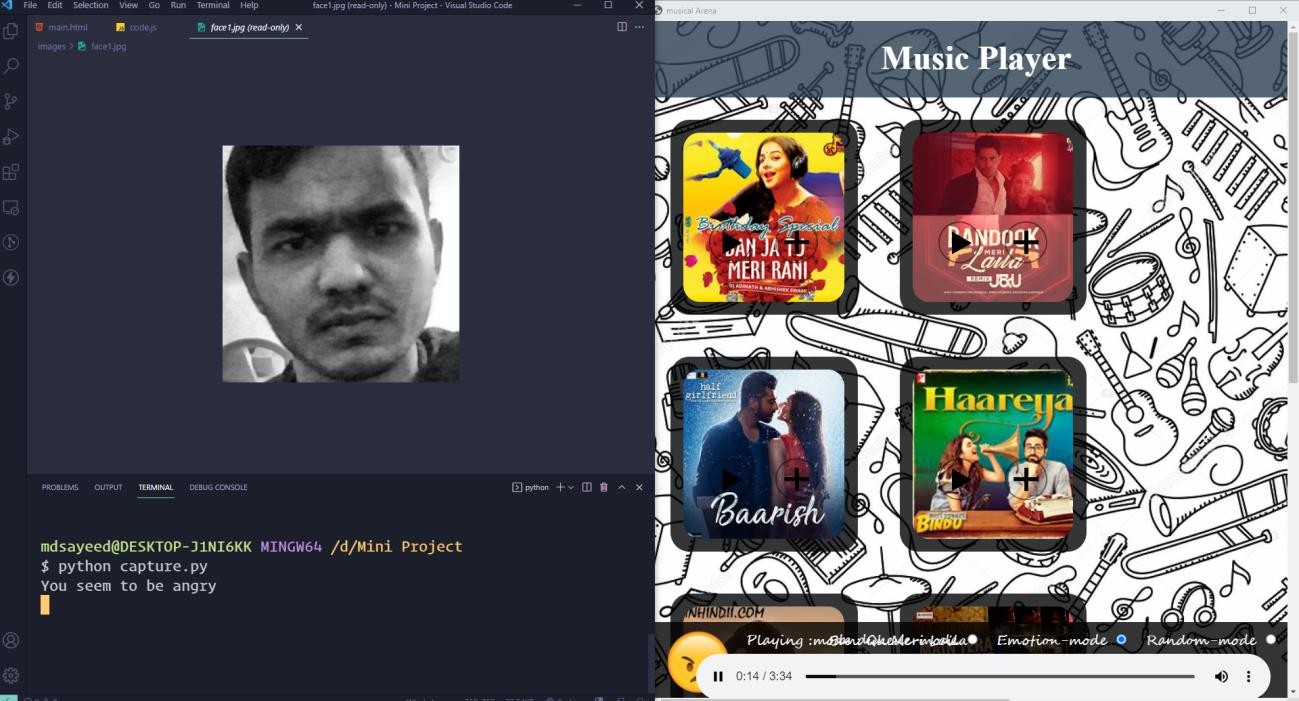
## Random Mode



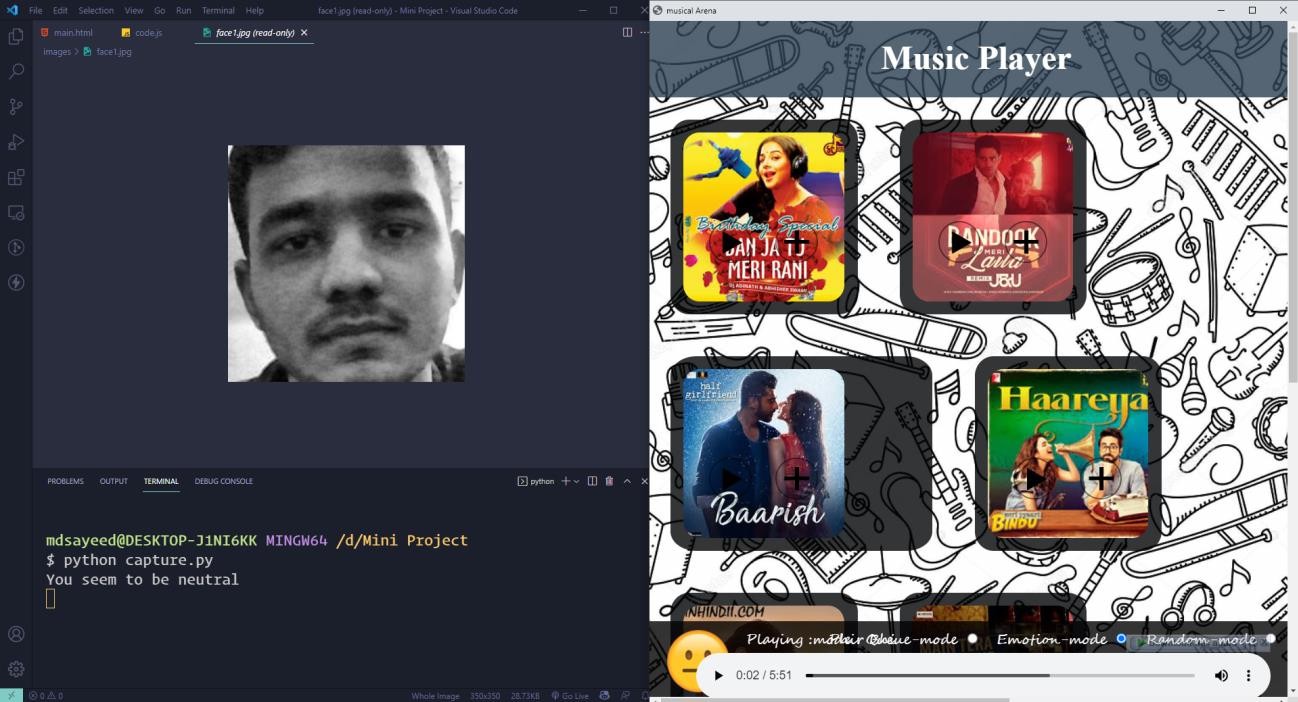
1. Emotion Mode Test case 1: Happy



Test case 2: Angry



Test case 3: Neutral



## Conclusion

Emotion-based music systems will be of great advantage to the users looking for music based on their mood and emotional behavior.

The system will help to reduce the time to search the music according to the mood of the user. By reducing the unnecessary time to compute, this increases the overall accuracy and efficiency of the system.

As a future work, we aim at upbringing a deployable, productive and theme-oriented project, thereby upskilling ourselves with the required tools and technologies.

## References

[1] Mandeep Kaur and Rajeev Vashisht, "Facial Expression Recognition using a Noval Approach and its Application", International Journal of Computer and Electrical Engineering, Vol. 3, No. 2, April 2011 (1793-8163).

[2] Maruthi Raja S K, Keerthi Vasan A, Kumaran V, Kavitha N, "Real Time Intelligent Emotional Music Player using Android “, Journal for Research | Volume 03| Issue 01 | March 2017 ISSN: 2395-7549.

[3] Faiyaz Mohammad Saif, Jabin Rubayat, Fahmida Sharmin Pranto, Md.Hosne Al Walid and Kazi Mohammed Shahanur Ahammed, "Facial Expressions Recognition Using Bezier Curve and Performance Analysis Using Ann, Naïve-Bayes and SVM Classifiers", Proceedings of 14th Global Engineering and Technology Conference 29-30 December 2017, BIAM Foundation, 63 Eskaton, Dhaka, Bangladesh ISBN: 978-1-925488-60-9.

[4] Tian, Y.L., Kanade, T. Cohn, J.F. (2005). Facial Expression Analysis, In Handbook of Face Recognition, Li, S.Z. Jain, A.K., (Eds.), pp. 247-276, Springer, New York, USA.

[5] P. M. Chavan, M. C. Jadhav, J. B. Mashruwala, A. K. Nehete, Pooja A. Panjari, “Real Time Emotion Recognition through Facial Expressions for Desktop Devices”, International Journal of Emerging Science and Engineering, Vol. 1, No. 7, May 2013.

[6] Yong-Hwan Lee Dankook University 152, Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do, 448-701, Korea, Detection and Recognition of Facial Emotion using Bezier Curves.

[7] M.S.Bartlett, G.Littlewort, I.Fasel, J.R.Movellan, Real-time face detection and expression Recognition: Development and application to human-computer interaction, Proceedings CVPR Workshop on computer vision and Pattern recognition for human-computer interaction.

[8] H.Rowley, S.Baluja, T.Kanade, and Neural Network based face detection, IEEE Trans.Pattern Analysis and Machine Intelligence, Vol.20, no.1pp 23-28.

[9] K.K.Sung & T.Poggio, Example-based learning for view-based human face detection, IEEE Transactions Pattern analysis and machine intelligence, Vol.20, No.1 pp: 39-51

[10] Prof. Jayshree Jha1, Akshay Mangaonkar2, Deep Mistry, Nipun Jambaulikar4, Prathamesh Kolhatkar5, "Facial Expression Based Music Player", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 10, October 2015

[11] Jyoti Rani, Kanwal Garg, "Emotion Detection Using Facial Expressions -A Review", Volume 4, Issue 4, April 2014 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering.

[12] Thayer ‖ the biopsychology of mood & arousal‖, Oxford University Press, 1989.

[13] Jia-Jun Wong, Siu-Yeung Cho, ‖ Facial emotion recognition by adaptive processing of tree structures‖.