Mood Spotifier with Chatbot

 A

Report submitted in partial fulfilment of the requirement for the

degree of

B.Tech.

In

###### Computer Science & Engineering

By

##### Hari Om Shukla (2101641520067)

##### Sarthak Jain(2101641520123)

##### Pratham Singh (2101641520108)

##### Alokik Prakash Gupta (2101641520012)

##### Under the guidance of

##### Dr. Subhash Singh Parihar

##### Associate Professor

##### Project Id: 22\_AI\_2A\_10



Pranveer Singh Institute of Technology, Kanpur

Dr A P J A K Technical University

Lucknow

## TABLE of CONTENTS

## DECLARATION……………………………………. iv

## CERTIFICATE……………………………………... v

## ACKNOWLEDGEMENTS………………………… vi

## ABSTRACT…………………………………………. vii

## LIST OF TABLES AND FIGURES.………………. viii

## CH. NO. TITLE PAGE NO.

## INTRODUCTION 9

## MOTIVATION 9

## BACKGROUND OF PROBLEM 9

## Current System 9

## Issues in Current System 9

## Functionality Issues 9

## Security Issues 10

## PROBLEM STATEMENT 10

## PROPOSED WORK 10

## ORGANIZATION OF REPORT 11

## DESIGN METHODOLOGY 13

## CHATBOT MODULE 13

## MOOD DETECTION MODULE 14

## Face Detection 14

## Mood Detection 14

## OpenCV 14

## MUSIC RECOMMENDATION MODULE 15

## INTEGRATION 16

## ii

## IMPLEMENTATION 17

## GATHERING DATA 17

## DATA PREPARATION 17

## Choosing a Model 17

## Using ML algorithm 17

## TRAINING 17

## EMOTION DETECTION USING FISHERFACE 18

## TESTING/RESULT AND ANALYSIS

## 4.1 RESULTS 20

## 4.2EXPERIMENT RESULTS- INSTRUCTIONSEXPLAINEDTO THE USER 20

## 4.3 EXPERIMENT RESULTS- INSTRUTIONS NOT EXPLAINED TO THE USER 20

## 4.4 RESIZING THE IMAGES 21

## CONCLUSION AND FUTURE ENHANCEMENTS 23

## LIMITATIONS 24

## FUTURE WORK 24

## REFERENCES………………………………………….. 38

## iii

## DECLARATION

***This is to certify that Report entitled “ MoodSpotifier with Chatbot” which is submitted by me in partial fulfilment of the requirement for the award of degree B.Tech. in Computer Science and Engineering to Pranveer Singh Institute of Technology, Kanpur Dr. A P J A K Technical University, Lucknow comprises only our own work and due acknowledgement has been made in the text to all other material used.***

#### Date:

***Hari Om Shukla(2101641520067)***

***Sarthak Jain(2101641520123)***

***Pratham Singh(2101641520108)***

***Alokik Prakash Gupta(2101641520012)***

**iv**

## CERTIFICATE

***This is to certify that Report entitled “Mood Spotifier with Chatbot” which is submitted by Hari Om Shukla, Sarthak Jain, Pratham Singh and Alokik Prakash Gupta in partial fulfilment of the requirement for the award of degree B.Tech. in Computer Science & Engineering to Pranveer Singh Institute of Technology, Kanpur affiliated to Dr. A P J A K Technical University, Lucknow is a record of the candidate own work carried out by him under my supervision. The matter embodied in this thesis is original and has not been submitted for the award of any other degree.***

***Signature: Signature:***

***Dr. Vishal Nagar S. S. Parihar***

***Dean CSE Department, Associate Professor***

***PSIT, Kanpur CSE Department,***

***PSIT, Kanpur***

## v

**ACKNOWLEDGEMENT**

***It gives us a great sense of pleasure to present the report of the B.Tech. Project undertaken during B.Tech. Second Year. We owe special debt of gratitude to our project supervisor S.S. Parihar , Department of Computer Science and Engineering, Pranveer Singh Institute of Technology, Kanpur for his constant support and guidance throughout the course of our work. His sincerely, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavours have seen light of the day.***

***We also take the opportunity to acknowledge the contribution of Professor Dr. Vishal Nagar, Dean Computer Science & Engineering Department, Pranveer Singh Institute of Technology, Kanpur for his full support and assistance during the development of the project.***

***We also do not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.***

***Signature Signature***

***Name: Hari Om Shukla Name:Sarthak Jain***

***Roll No.:2101641520067 Roll No.:2101641520123***

***Signature Signature***

***Name: Pratham Singh Name:Alokik Prakash Gupta***

***Roll No.:2101641520108 Roll No.:2101641520012***

**vi**

**ABSTRACT**

**It is often confusing for a person to decide which music he/she have to listen from a massive collection of existing options. There have been several suggestion frameworks available for issues like music, dining, and shopping depending upon the mood of user. The main objective of our music recommendation system is to provide suggestions to the users that fit the user’s preferences. The analysis of the facial expression/user emotion may lead to understanding the current emotional or mental state of the user. Music and videos are one region where there is a significant chance to prescribe abundant choices to clients in light of their inclinations and also recorded information. It is well known that humans make use of facial expressions to express more clearly what they want to say and the context in which they meant their words. More than 60 percent of the users believe that at a certain point of time the number of songs present in their songs library is so large that they are unable to figure out the song which they have to play.**

**By developing a recommendation system, it could assist a user to make a decision regarding which music one should listen to helping the user to reduce his/her stress levels. The user would not have to waste any time in searching or to look up for songs and the best track matching the user’s mood is detected, and songs would be shown to the user according to his/her mood. The image of the user is captured with the help of a webcam. The user’s picture is taken and then as per the mood/emotion of the user an appropriate song from the playlist of the user is shown matching the user’s requirement.**

**This project is a novel approach that helps the user to automatically play songs based on the emotions of the user. It recognizes the facial emotions of the user and plays the songs according to their emotion. The emotions are recognized using a machine learning method Support Vector Machine (SVM )algorithm. SVM can be used for classification or regression problems. It finds an optimal boundary between the possible outputs. The training dataset which we used is Olivetti faces which contain 400 faces and its desired values or parameters. The webcam captures the image of the user.It then extract the facial features of the user from the captured image. The training process involves initializing some random values for say smiling and not smiling of our model, predict the output with those values, then compare it with the model's prediction and then adjust the values so that they match the predictions that were made previously. Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world. According to the emotion ,the music will be played from the predefined directories Key Words: SVM, Olivetti faces, Emotions, Songs, Machine learning, Training, Testing.Listening to music affects the human brain activities. Emotion based music player with automated playlist can help users to maintain a particular emotional state. vii** **LIST OF TABLES AND FIGURES**

**TABLES:-**

**Table 1 Instructions explained to the user 20**

**Table 2 Instructions not explained to the user 21**

**FIGURES:-**

**Figure 1 Architecture Diagram of Chatbot 13**

**Figure 2 ER Diagram of the Project 14**

**Figure 3 Architecture of Music Recommender 15**

**Figure 4 DFD of the Project 16**

**Figure 5 Training Perfect Images 18**

**Figure 6 Inheritance of Fisher-Face Recognizer 19**

**Figure 7 Experiment Results-Instructions explained to user 20**

**Figure 8 Experiment Results-Instructions not explained to user 21**

**Figure 9 Convolution Neural Network 22**

**Figure 10 Training and Validation Accuracy & Loss 24**

**viii**

## CHAPTER 1

## INTRODUCTION

## MOTIVATION

## People tend to express their emotions, mainly by their facial expressions. Music has always been known to alter the mood of an individual.

## Capturing and recognizing the emotion being voiced by a person and displaying appropriate songs matching the one's mood and can increasingly calm the mind of a user and overall end up giving a pleasing effect.

## BACKGROUND OF PROBLEM

## Current System

## The music players which are developed in previous days have very less features compared to this newly build music player. It has both the old school features and latest emotion-based auto music playing feature. Most of the music players available only have manual selection of the playlist and random song shuffle.

## Besides that, this music player can judge user's mood and using previous data it creates the playlist.

## Currently, there are many existing music player applications. Some of the interesting applications among them are: ● Saavan and Spotify { These application [16] gives good user accessibility features to play songs and recommends user with other songs of similar genre. ● Moodfuse - In this application [12], user should manually enter mood and genre that wants to be heard and moodfuse recommends the songs-list. ● Steromood - User should select his mood manually by selecting the moods from the list and the application [17] plays music from YouTube.

## Issues in Current Sytem

## Functionality Issues

**They can’t do well in different databases.**

**Only few songs due to less storage space available.**

**9**

**Unable to identify complex and mixed emotions. The features available uses only a single monochrome intensity image.Unable to judge only one mood at a time. Future Work could be done to solve problems:**

## Security Issues

**Use of Convolutional Neural Network (CNN) to solve database problem.**

**In case of storage problem, the songs can be exported to cloud storage and users will have the option of downloading songs as of their choice.**

**Train to individually learn the features for recognizing different objects and categories.**

* 1. **PROBLEM STATEMENT**

**Using traditional music players, a user had to manually browse through his playlist and select songs that would soothe his mood and emotional experience.**

**In today‘s world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams and including volume modulation, genre classification etc.**

**Although these features satisfy the user‘s basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour. That is the requirements of an individual, a user sporadically suffered through the need and desire of browsing through his playlist, according to his mood and emotions**.

* 1. **PROPOSED WORK**

**The proposed system can capture the user's facial expressions, and based on his facial expressions, extract facial reference points, and then classify them to obtain the user's specific emotions. After the emotions are classified, songs corresponding to the user's emotions are displayed to the user.Facial expressions tell us about the mood and gives us clue to know a person’s emotion. People tends to express their emotions through facial expressions. Music is one thing that can alter a person’s mood. Lips and Eyes do not just have sensory receptors, they also show us in what mood that person is in. 10**

**The work describes a computer application in the form of “Emotion based Smart Music Player” which provides the users an added ease in creating playlist and playing music. The project aims to capture a person’s emotions through facial expressions and lighten the mood of the user, by playing song which suits to that situation.**

**It captures and recognizes the emotion that a person is expressing, and can gradually calm the user’s mind, so it usually has a pleasant effect. The music player is designed to capture human emotions with the help of webcam interface available in the computer system. When the application starts, the system captures the user’s picture. The picture was taken via webcam. The previously captured image will be saved and rendering phase will start. After a period, the user’s mood may change, and it may or may not change. Therefore, the image is captured after every song or decided interval of time. And the image will be forwarded to the next phase. In some cases, mood alteration may also help in overcoming situations like depression and sadness. Expression analysis can avoid many health risks and take measures to improve the user’s mood.**

**The image recognition system is classified into two types, feature-based system, and image-based system. In the initial system, options extracted from the image parts, nose, mouth, lips, etc. That are then sculpturesque to confirm the relation between these options. Whereas in the second system, image pixels are used and represented as inbound methods.In the model few sample models are included such as Happy, Sad, Angry, and Neutral. It also has mixed mood feature. Every sample model is assigned with some songs according to the classification of user. Emotion based music players available now in use are more time consuming than our proposed system.**

## ORGANIZATION OF REPORT

**The main concept of this project is to automatically play songs based on the emotions of the user. It aims to provide user-preferred music with emotion awareness. In existing system user want to manually select the songs, randomly played songs may not match to the mood of the user, user has to classify the songs into various emotions and then for playing the songs user has to manually select a particular emotion. These difficulties can be avoided by using Emo Player (Emotion based music player). The emotions are recognized using a machine learning method Support Vector Machine (SVM)algorithm. SVM can be used for classification or regression problems. According to the emotion, the music will be played from the predefined directories. In some cases, mood alteration may also help in overcoming situations like depression and sadness. With the aid of expression analysis, many health risks can be avoided, and also there can be steps taken that help brings the mood of a user to a betterstage.**

**11**

**Expressing and recognizing emotions of human are very much important in communication system [3].**

**Human beings have the ability to express and recognize emotions. Computer seeks to identify the human emotions either by image analysis or through sensors [6]. In our day to daylife and in our professional life we interact with many people face to face or indirectly by phone calls, sometimes it is necessary for people to be aware of their present emotions of the person with whom they are interacting. Human emotions are classified as: surprise, fear, anger, happy, sad, disgust and neutral [3].**

**Facial movement [1] and the tone of speech play a major role in expressing emotions. The physique and tone of the face tells the energy in the utterance of speech, which can be firstly modified to communicate different feelings. Humans can easily recognize these changes in signals along with the information felt by any other sensory organs. This project analyses the use of image or sensors or speech to capture the emotionsMusic plays [4] a vital role in enhancing an individual’s life as it is an important medium of entertainment for music lovers and listeners [15] and sometimes even imparts a therapeutic approach. Where words fail music speaks‖, and hence it can change person’s negative emotion simultaneously and slowly into a positive mood [7].**

**Emotions can be expressed through gestures, speech, facial expressions, bodylanguage etc. For the system to understand the user’s mood, we use facial expression [3]. Using the mobile 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. For this application, we are using neural networks for recognition of emotion[9][8].**

**At times it is possible that user might like different kinds of songs in certain mood. For example, when a user’s emotion is detected to be Sad, then it is totally users choice what kind of mood does he/she wants. There are two possibilities in this scenario: a) User wants to continue his/her sad mood. b) User wants to elevate his/her mood and wants to be happy. Therefore, depending on the choice of users the songs in the sub directories can be changed. As the program runs successfully on system.**

**People tend to express their emotions, mainly by their facial expressions. Music has always been known to alter the mood of an individual. Capturing and recognizing the emotion being voiced by a person and displaying appropriate songs matching the one's mood and can increasingly calm the mind of a user and overall end up giving a pleasing effect. The project aims to capture the emotion expressed by a person through facial expressions. A music player is designed to capture human emotion through the web camera interface available on computing systems.**

**It captures and recognizes the emotion that a person is expressing, and can gradually calm the user’s mind, so it usually has a pleasant effect. The music player is designed to capture human emotions with the help of webcam interface available in the computer system. When the application starts, the system captures the user’s picture. The picture was taken via webcam. The previously captured image will be saved and rendering phase will start. After a period, the user’s mood may change, and it may or may not change. Therefore, the image is captured after every song or decided interval of time. And the image will be forwarded to the next phase. In some cases, mood alteration may also help in overcoming situations like depression and sadness. 12**

## CHAPTER 2

## DESIGN METHODOLOGY

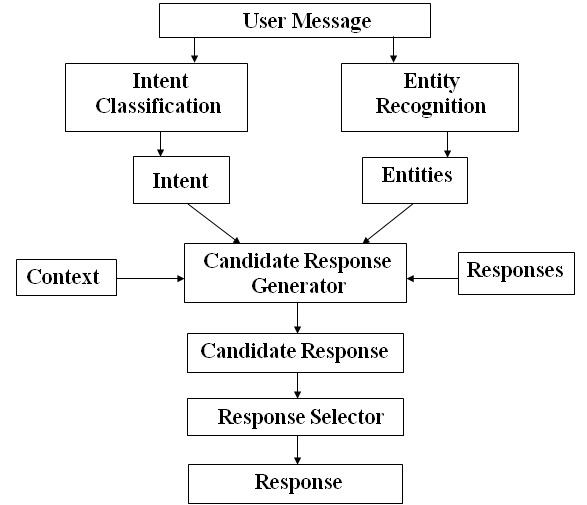
**2.1 Chatbot Module**

**This system starts with a conversation to the user like “Hello”, ”What’s Your Mood”, ”Let me check it”, etc. It is also used for getting feedback from the user. Chatbot will be implemented in python language using Chatterbot Library.**

**ChatterBot is a Python library designed to make it easy to create software that can engage in conversation. An untrained instance of ChatterBot starts off with no knowledge of how to communicate. Each time a user enters a statement, the library saves the text that they entered and the text that the statement was in response to.**

**As ChatterBot receives more input the number of responses that it can reply and the accuracy of each response in relation to the input statement increase. The program selects the closest matching response by searching for the closest matching known statement that matches the input, it then chooses a response from the selection of known responses to that statement.**

**The mood-based music recommendation system is an application that focuses on implementing real time mood detection. It is a prototype of a new product that comprises two main modules: Facial expression recognition/mood detection and Music recommendation**

****

**Figure 1. Architecture Diagram of Chatbot 13**

**2.2 Mood Detection Module**

**This Module is divided into two parts:**

**2.2.1 Face Detection**

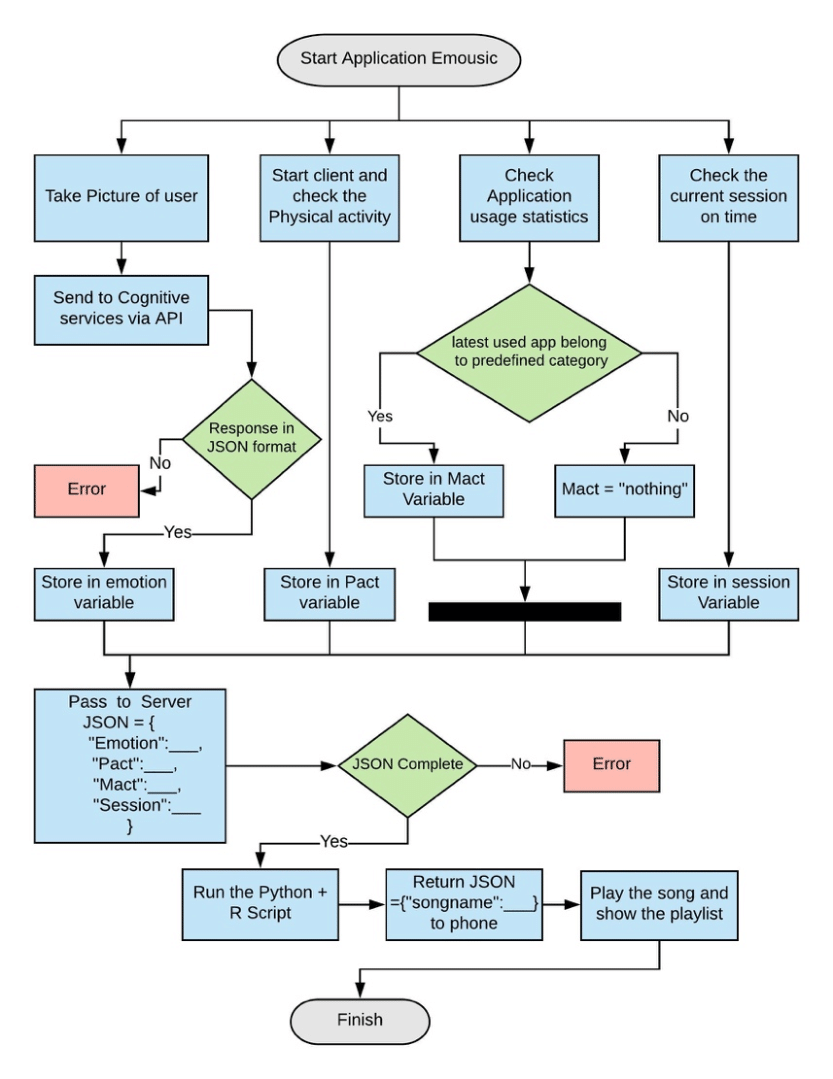
**Ability to detect the location of face in any input image or frame. The output is the bounding box coordinates of the detected faces. For this task, the python library OpenCV was considered.**

**2.2.2 Mood Detection**

**Classification of the emotion on the face as happy, angry, sad, neutral, surprise, fear or disgust. Keras which is a CNN architecture model for Image Classification and Mobile Vision was used. There are other models as well but what makes Keras special is that it has very less computation power to run or apply transfer learning to. This makes it a perfect fit for Mobile devices, embedded systems and computers without GPU or low computational efficiency without compromising the accuracy of the results. It uses depth wise separable convolutions to build light weight deep neural networks. The dataset used for training was obtained by combining FER 2013 dataset [6] and MMA Facial Expression Recognition dataset [7] from Kaggle. The FER 2013 dataset contained grayscale images of size 48x48 pixels. The MMA Facial Expression Recognition dataset had images of different specifications. Thus, all these images were converted as per the images in FER 2013 dataset and combined to obtain an even larger dataset with 40,045 training images and 11,924 testing images. Keras was used to train and test our model for seven classes - happy, angry, neutral, sad, surprise, fear and disgust. We trained it for 25 epochs and achieved an accuracy of approximately 75%.**

**2.2.3 OpenCV**

**Open CV provides many functions for facial detection and facial recognition. It comes with a trainer and a detector. If you want to train your own classified objects such as mobile phones, pens, etc... you can use Open CV to create it.**

****

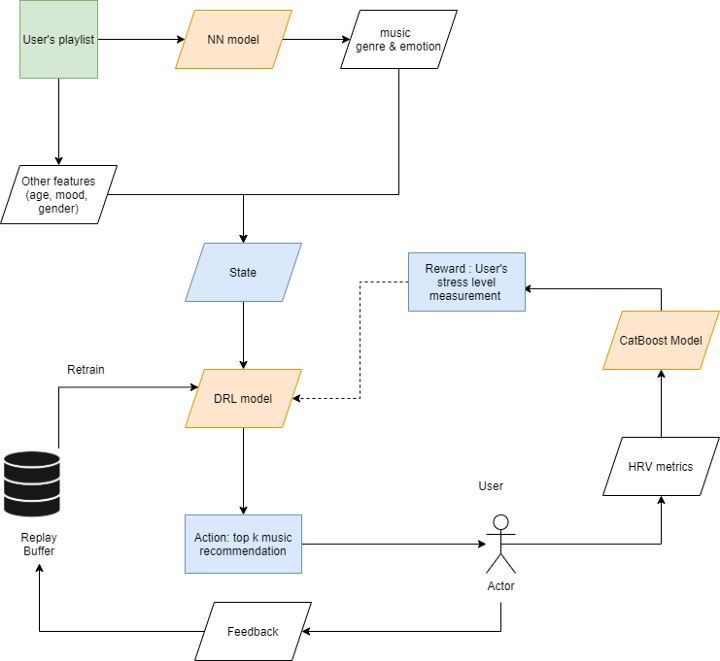
**Figure 2. ER Diagram of the Project**

**14**

**2.3 Music Recommendation Module**

**Once you know exactly what you want and the equipment’s are in hand, it takes you to the first real step of machine learning- Gathering Data. This step is very crucial as the quality and quantity of data gathered will directly determine how good the predictive model will turn out to be. The data collected is then tabulated and called as Training Data.**

**After the training data is gathered, you move on to the next step of machine learning: Data preparation, where the data is loaded into a suitable place and then prepared for use in machine learning training. Here, the data is first put all together and then the order is randomized as the order of data should not affect what is learned. This is also a good enough time to do any visualizations of the data, as that will help you see if there are any relevant relationships between the different variables, how you can take their advantage and as well as show you if there are any data imbalances present. Also, the data now has to be split into two parts. The first part that is used in training our model, will be the majority of the dataset and the second will be used for the evaluation of the trained model’s performance. The other forms of adjusting and manipulation like normalization, error correction, and more take place at this step.**

**.**

**Figure 3. Architecture of Music Recommender**

**15**

**2.4 Integration**

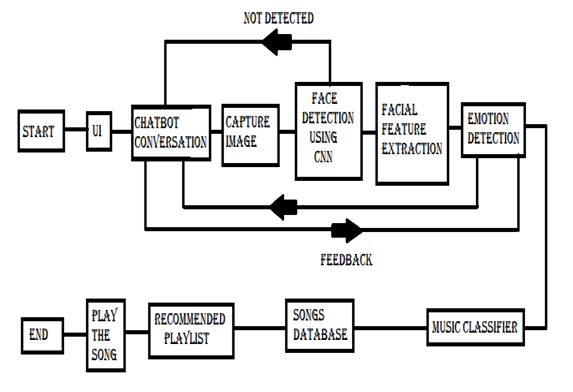
**For the integration of these two modules in an Android application, the trained Keras model was saved as an .h5 file, and this .h5 file was then converted to a .tflite file using TensorFlow Lite Converter. It takes a TensorFlow model as input and generates a TensorFlow Lite model as output with .tflite extension. Since the Keras model is used, the size of the tflite file is expected be around 20- 25 Megabyte (MB) which was the desired size. The labels.txt file contains the class labels of the model. All the appropriate methods were created for loading the model, running the interpreter and obtaining the results.**

**All the photos present in dataset are firstly converted to grayscale, for making preprocessing and detection more efficiently, faster and easier. Each input image is in form of pixels (e.g. 48x48). Now the pixel represented images are sent to the convolution layers (hidden layers). In Between each layer maximum pooling is done, the purpose of doing so, is to down-sample the input data or image, reducing the dimensions and allows assumption to be made about features contained in sub regions.**

**This is done to avoid over-fitting.**

**As well as it reduces computational cost by reducing number of parameters to learn. Example, if input image is of matrix 4x4 representation and let’s say output we want is in 2x2, then pooling is performed in between all hidden layers. After that data is sent to dense layer, to prevent over-fitting.**

**Dropout technique is used to reduce over-fitting in neural networks. The output layer conveys the detected class. Let’s say if the detected expression is happy, then the next step is to select anyone training dataset for music model. Now, the dataset is trained according to the match for playing music. LSTM neural network is used for classifying the songs. One hot encoding is performed to represent categorical variables into binary vectors, so as to make the classification faster and better.**

****

## Figure 4. DFD of the Project

## 16

## CHAPTER 3

## IMPLEMENTATION

## 3.1Gathering Data

## Once you know exactly what you want and the equipment’s are in hand, it takes you to the first real step of machine learning- Gathering Data. This step is very crucial as the quality and quantity of data gathered will directly determine how good the predictive model will turn out to be. The data collected is then tabulated and called as Training Data.

**3.2 Data Preparation**

## After the training data is gathered, you move on to the next step of machine learning: Data preparation, where the data is loaded into a suitable place and then prepared for use in machine learning training. Here, the data is first put all together and then the order is randomized as the order of data should not affect what is learned. This is also a good enough time to do any visualizations of the data, as that will help you see if there are any relevant relationships between the different variables, how you can take their advantage and as well as show you if there are any data imbalances present. Also, the data now has to be split into two parts. The first part that is used in training our model, will be the majority of the dataset and the second will be used for the evaluation of the trained model’s performance. The other forms of adjusting and manipulation like normalization, error correction, and more take place at this step.

## 3.2.1 Choosing a model

## The next step that follows in the workflow is choosing a model among the many that researchers and data scientists have created over the years. Make the choice of the right one that should get the job done.

## 3.2.2 Using ML algorithms

## Machine learning is a field of computer science that gives computers the ability to learn without being programmed explicitly. The power of machine learning is that you can determine how to differentiate using models, rather than using human judgment.

## 3.3 Training

## After the before steps are completed, you then move onto what is often considered the bulk of machine learning called training where the data is used to incrementally improve the model’s ability to predict. The training process involves initializing some random values for say A and B of our model, predict the output with those values, then compare it with the model's prediction and then adjust the values so that they match the predictions that were made

## 17

## previously. This process then repeats and each cycle of updating is called one training step.

## It detects feature points on its own. For facial recognition, the RGB image is first converted into a binary image. If the average pixel value is less than 110, black pixels are used as substitute pixels, otherwise, white pixels are used as substitute pixels.

## Now, every possible models and positions of all cores are used to estimate many functions. But of all these functions we calculated, most of them are not relevant. The figure below shows two good attributes in the first row. The first function selected seems to focus on the attribute that the eye area is usually darker than the nose and cheek areas. The second function selected is based on the fact that the eyes are darker than the bridge of the noseScreenshot (270).png

## Figure 5. Training Perfect Images

## 3.4 Emotion Detection using Fisherface Method

## Fisherface is one of the simplified and most preferred algorithms used in facial and emotion recognition.

## It is a higher-level and better technique because it works and performs great in maximizing the difference between the classes in the training process.

## Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) methods are used to reduce the face space dimension.

## Minimum Euclidean is calculated and used to identify the matching image in fisherfaces algorithm. Fisherface Method works as a merger between PCA and LDA methods.

## LDA analyse and makes the prediction by calculating the probability from the set of inputs given to every class.

## After estimating the class with the highest probability is the output class and the prediction is made.

## PCA is most used technique for dimension reduction, it finds all the group of variables with maximum variance possible using orthogonal transformation. 18

## Once training is complete, you now check if it is good enough using this step. This is where that dataset you set aside earlier comes into play.

## Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world.

## Parameter Tuning: Once the evaluation is over, any further improvement in your training can be possible by tuning the parameters. There were a few parameters that were implicitly assumed when the training was done.

## Another parameter included is the learning rate that defines how far the line is shifted during each step, based on the information from the previous training step.

## These values all play a role in the accuracy of the training model, and how long the training will take. For models that are more complex, initial conditions play a significant role in the determination of the outcome of training.

## Differences can be seen depending on whether a model starts off training with values initialized to zeroes versus some distribution of values, which then leads to the question of which distribution is to be used.

## Since there are many considerations at this phase of training, it’s important that you define what makes a model good. These parameters are referred to as Hyper parameters.

## The adjustment or tuning of these parameters depends on the dataset, model, and the training process. Once you are done with these parameters and are satisfied you can move on to the last step.

## Screenshot (275).png

## Figure 6. Inheritance of Fisher-Face Recognizer

## 19

## CHAPTER 4

## TESTING/RESULT AND ANALYSIS

## 4.1 Results

## This study proposes a music recommendation system which extracts the image of the user, which is captured with the help of a camera attached to the computing platform. Once the picture has been

## 4.2Experiment Results-Instructions Explained to the User:

## In this scenario the users were given instructions as to what is to be done to perform the prediction of the emotion expressed which provided the following results.

## Sometimes in cases where the inner emotion is sad and facial expression is happy it resulted in a fail case. The values are given in Table 1 and the result is shown in Figure 2.

## Screenshot (271).png

## Table 1. Instructions explained to the User

## Screenshot (272).png

## Figure 7. Experiment Results-Instructions explained to User

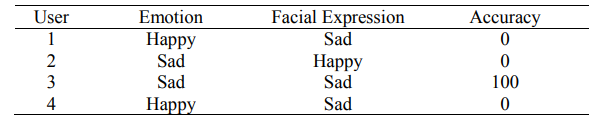
## 4.3 Experiment Results-Instructions not given to the User:

## In this scenario the users were not given any instructions as to what is to be do and thus the inner emotions or the emotions recognized failed, there were also cases where in the emotion

## matched with the facial expressions of the user. 20

## The values are given in Table 2 and the result is shown in Figure 3.After initialising the camera, according to the user input it captures the images and sent for analysing. For every success capturing of image the new image will be created.

## If the analysing of captured image sets fails, then again, a new image will be captured and it continues until it captures a perfect image which can be analysed further (as shown in figure 3.1).



## Table 2.Instructions not explained to the User

## Screenshot (273).png

## Figure 8. Experiment Results-Instructions not explained to User

## 4.4 Resizing the images

## Whatever image we have got chosen for the dataset it principally involving the dimensions which may offer a precise output. The size is chosen such that the model can able to simply distinguish face from image by haarcascade model.

## And therefore, the size we tend to get from real-time-scan is not always same as data, but it has very less difference. In our model it takes 350\*350 as size of the image. Gray Scaling of images:

## This is the requirement of the model because it helps in acquiring better results. The shaded face and contrast can give us reliable results. As RGB are three dimensional while grayscale images are single dimensional it helps as Dimension Reduction. And it also reduces model complexity.

## Euclidean algorithm formula (figure 3.2) is used to find the distance of difference between the image testing with training face image. 21

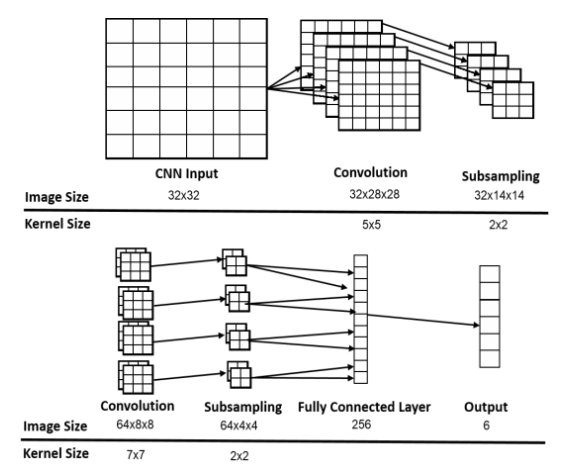
## The result will be the image that has the smallest distance with the test mage displayed by the system.

## Fisher face method is used for analysing the extracted image and performs training. It follows a three-step process for training the perfect images (as shown in figure 3.3).

## Those extracted features are used in mood detection module to detect the mood of the user and it passes the result to the backend of music player.

## With the results obtained it checks for the lists which are assigned to each mood, after checking the mood it plays the songs available in that list.

## User can also change the list of his favourite songs. Some sample moods available are Happy, Sad, Angry, Neutral.

****

**Figure 9. Convolution Neural Network**

**22**

## CHAPTER 5

## CONCLUSION AND FUTURE ENHANCEMENTS

## The proposed system processes images of facial expressions, recognizes the actions related to basic emotions, and then plays music based on these emotions. The main advantage of this system is that it is completely independent of automation.

## In the future, the application can export songs to a dedicated cloud database and allows users to download desired songs, as well as to recognize complex and mixed emotions.

## Therefore, the developed application will provide users with the most suitable songs based on their current emotions, thereby reducing the workload of users creating and managing playlists, bringing more fun to music listeners, not only helping users, but also songs can be organized systematically.

## And this model also has productive future scope, facial recognition can be used for authentication purpose, an android app can be developed for mobile uses, it can also detect sleepy mood if the driver is not focused on driving.

## Even people who are physically challenged can use this, for them this is very better than voice-based applications. Emotion recognition using facial expressions is one of the important topics of research and has gathered much attention in the past.

## It can be seen that the problem of emotion recognition with the help of image processing algorithms has been increasing day by day. Researchers are continuously working on ways to resolve this by the use of different kinds of features and image processing methods.

## The applications of image processing algorithms in the field of both medical science and human science are of vast importance.

## There are continuously new ways and methods being developed that make use of image processing algorithms to extract the emotion of the user and make use of the extracted emotion to treat the user.

## Emotion recognition has gained a lot of importance in all aspects of life and if a robust algorithm implemented which can accurately classify the emotions of the person, then a great deal of advancement in the industry can be achieved with the help of this.

## The system has successfully been able to capture the emotion of a user.

## It has been tested in a real-time environment for this predicate.

## However, it has to be tested in different lighting conditions to determine the robustness of the developed system. The system has also been able to grab the new images of the user and appropriately update its classifier and training dataset.

## The system was designed using the facial landmarks scheme and was tested under variousscenarios for the result that would be obtained.

## 23

## It is seen that the classifier has an accuracy of more than 80 percent for most of the test cases, which is pretty good accuracy in terms of emotion classification.

## It can also be seen that the classifier can accurately predict the expression of the user in a real-time scenario when tested live for a user.

## 5.1 Limitations

## The system still is not able to record all the emotions correctly due to the less availability of the images in the image dataset being used. The image that is fed into the classifier should be taken in a well-lit atmosphere for the classifier to give accurate results. The quality of the image should be at least higher than 320p for the classifier to predict the emotion of the user accurately.

## 5.2 Future Work

## Reduce the time required to train the classifier . Use of EEG signals to make the software even more optimized and to detect the exact mood /emotion of the user.

## In future Music Player can be enhanced with Google play music, so songs which are not present in local storage can also be played and to access the whole application in speech based. The Emotion Based Music System will be of great advantage to users looking for music based on their mood and emotional behavior. It will help reduce the searching time for music thereby reducing unnecessary time and hence increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the musictherapist to treat the patient. In future it can also be used to detect the sleepy mood of the driver, driving the car and many more uses. Also with its additional features mentioned above, it will be a complete system for music lovers and listeners.

## model_training_results_1.jfifmodel_training_results_2.jfif

## 

## Figure 10. Training and Validation Accuracy & Loss

## 24

## Screenshot (276).pngScreenshot (277).pngScreenshot (278).pngScreenshot (279).pngScreenshot (280).pngScreenshot (281).pngScreenshot (282).pngScreenshot (283).pngScreenshot (284).pngScreenshot (285).pngScreenshot (286).pngScreenshot (287).pngScreenshot (288).png

**25**

**26**

**27**

**28**

**29**

**30**

**31**

**32**

**33**

**34**

**25**

**35**

**36**

**37**

**REFERENCES**

* **Geeks for Geeks -** [**https://www.geeksforgeeks.org**](https://www.geeksforgeeks.org)
* **Ijert -** [**https://www.ijert.org/research/mood-based-music-recommendation-system-IJERTV10IS060253.pdf**](https://www.ijert.org/research/mood-based-music-recommendation-system-IJERTV10IS060253.pdf)
* **Towardsdatascience -** [**https://towardsdatascience.com/build-your-first-mood-based-music-recommendation-system-in-python-26a427308d96**](https://towardsdatascience.com/build-your-first-mood-based-music-recommendation-system-in-python-26a427308d96)
* **Buildmedia -** [**https://buildmedia.readthedocs.org**](https://buildmedia.readthedocs.org)

# Applied DL & CV for Images & Video Data Analysis from Eduxlabs

* [**https://www.eduxlabs.com/live-dl-cv-training**](https://www.eduxlabs.com/live-dl-cv-training)
* **Edureka! -** [**https://www.youtube.com/watch?v=G1Uhs6NVi-M&t=565s**](https://www.youtube.com/watch?v=G1Uhs6NVi-M&t=565s)
* **Intellipaat -** [**https://www.youtube.com/watch?v=FV3IvHeuH\_k&t=364s**](https://www.youtube.com/watch?v=FV3IvHeuH_k&t=364s)
* **Codebasics -** [**https://www.youtube.com/watch?v=zfiSAzpy9NM**](https://www.youtube.com/watch?v=zfiSAzpy9NM)
* **CodeWithHarry -** [**https://www.youtube.com/watch?v=RhEjmHeDNoA**](https://www.youtube.com/watch?v=RhEjmHeDNoA)
* **Iopscience -**[**https://iopscience.iop.org/article/10.1088/1757-899X/912/6/062007/pdf**](https://iopscience.iop.org/article/10.1088/1757-899X/912/6/062007/pdf)
* **Core -** [**https://core.ac.uk/download/pdf/211922219.pdf**](https://core.ac.uk/download/pdf/211922219.pdf)

**38**