

# Emotion Based Music Recommendation System

Sriraj Katkuri<sup>1</sup>, Mahitha Chegoor<sup>2</sup>, Dr.K.C. Sreedhar<sup>3</sup>, M. Sathyanarayana<sup>4</sup>

<sup>3,4</sup> Professor, Dept. of Computer Science and Engineering, SNIST, Hyderabad-501301, India

<sup>1,2</sup> B. Tech Scholar, Dept. of Computer Science and Engineering, SNIST Hyderabad-501301, India

**Abstract:-** Emotion-based music recommendation is a challenging task that has attracted significant research attention in recent years. This paper proposes a novel approach to emotion-based music recommendation using Support Vector Machines (SVMs). Our approach uses images as inputs instead of pronouns to capture the emotional state of the user more accurately. We collected a dataset of images associated with different emotional states and extracted features using Convolutional Neural Networks (CNNs). Then, we trained an SVM model to predict the emotional state of the user based on the input image. Finally, we used the output of the SVM model to recommend music tracks that are most suitable for the user's emotional state. Our experiments on a large dataset show that our approach outperforms existing methods in terms of accuracy and user satisfaction. The proposed approach provides a promising direction for emotion-based music recommendation systems that can enhance the user experience by providing personalized music recommendations based on their emotional state.

**Keywords:-** Haar Cascade; Convolutional Neural Networks; Recurrent Neural Network; Emotion Classification; Inception V3; MobileNet; ResNet-50

## I. INTRODUCTION

### 1.1 MOTIVATION

Our objective for this study is to utilize feeling ID procedures connected with wearable figuring devices to make extra commitments for the music proposition system's estimation, thus chipping away at the accuracy of the accompanying music thoughts.

The review or practice of making, building, utilizing, or concentrating on body-worn computational and tactile gadgets that empower a clever sort of human-PC cooperation with a body-connected, consistently on part is known as wearable figuring. As the amount of people using wearable handling contraptions creates numerous years, so do their applications. Clinical, wellness, inability, schooling, transportation, monetary, gaming, and music ventures have all been affected by them. Idea engines are estimations that channel critical information from an enormous pool of data to give the main things to the client. By learning client inclinations and delivering results that are relevant to their necessities and interests, suggestion motors can possibly uncover information designs in an assortment.

### 1.2 PROBLEM DEFINITION

In prior work, we analyzed feeling affirmation using only GSR signals. In this work, we add PPG to signs and deal an information combination based strategy for recognizing feelings in music proposal motors. The proposed wearable related music proposition designing considers the client's economics, yet furthermore their significant condition right now of thought. Using GSR and PPG information, we found empowering results for feeling expectation.

### 1.3 OBJECTIVE OF THE PROJECT

Most of the proposed frameworks don't consider human sentiments or feelings. Sentiments, on the other hand, influence people's ordinary everyday presences. For various applications, including human-robot connection, PC helped coaching, feeling mindful intuitive games, neuromarketing, and socially insightful programming applications, PCs ought to examine the feelings of their human discussion accomplices. Feelings can be recognized utilizing discourse investigation and looks. Nevertheless, by virtue of individuals who wish to cover their opinions, using simply talk or look signs may not be adequate to constantly recognize sentiments. It is more solid to screen and perceive individuals' inward mental cycles and feelings by using physiological signs instead of visual articulations.

## II. LITERATURE SURVEY

### 2.1 EXISTING SYSTEM:

Before as of late, most of PDS research zeroed in on the best way to defend information put away in the PDS and uphold client protection settings. Then again, the vital issue of helping clients in choosing their security settings for PDS information has gotten less consideration. Most of PDS clients don't have any idea how to make an interpretation of their security necessities into a bunch of protection settings, so this is a significant issue. As indicated by various examinations, normal clients may not precisely depict possibly muddled protection choices.

### DISADVANTAGES OF EXISTING SYSTEM:

Carefully created individual data is scattered across numerous web-based frameworks oversaw by different specialist co-ops (like internet based virtual entertainment, medical clinics, banks, aircrafts, etc). Clients let completely go over their

information, whose upkeep is the obligation of the information supplier, and can't completely use their information on the grounds that every supplier keeps an unmistakable picture of them.

## 2.2 PROPOSED SYSTEM:

Personal Data Storage (PDS) has denoted a critical change in how people can store and deal with their own information by changing from a help driven to a client driven approach. PDSs can be utilized by people to solidify their own information into a solitary sensible vault. From that point forward, these sorts of information can be associated and utilized by the right logical instruments, and they can likewise be imparted to others heavily influenced by the end client.

## ADVANTAGES OF THE PROPOSED SYSTEM:

We gave a technique to grouping music as happy, sad, angry, and different sorts. Feeling Based-Music-Player is a music player with Chrome as the front end that utilizes a Python-based ML calculation to perceive feelings in the client's face.

## III. BACKGROUND STUDY

The use of technology to detect human emotions has been a topic of interest in various fields, including psychology, neuroscience, and computer science. Facial expressions are considered one of the most reliable and valid indicators of human emotions. Therefore, the detection of facial expressions using machine learning algorithms has been widely studied in recent years.

The Haar Cascade Classifier is a popular machine learning technique for detecting facial expressions. Viola and Jones published this technique in 2001, and it is based on a machine learning-based object identification approach. The Haar Cascade Classifier trains a classifier to recognise objects in fresh photos using a set of positive and negative training images.

In recent years, the Haar Cascade Classifier has been widely used in various applications, including facial expression detection. The algorithm has been used to detect emotions such as happiness, sorrow, anger, surprise, fear, and contempt.

The use of facial expression detection in personalized music recommendation is a novel approach that has gained significant attention in recent years. The idea behind this approach is that music can influence human emotions, and emotions can be detected from facial expressions. Therefore, by detecting the emotions from facial expressions, music can be recommended based on the detected emotion.

Several studies have been conducted to investigate the use of facial expression detection in personalized music recommendations. For instance, a study conducted by Yang et al. (2018) used facial expression detection to recommend music based on the detected emotion. The study found that the proposed approach outperformed the traditional music recommendation systems.

In conclusion, the use of machine learning algorithms, such as the Haar Cascade Classifier, for facial expression detection has been widely studied in recent years. The proposed emotion-based music recommendation system, which uses facial expression detection to recommend music based on the detected emotion, has potential applications in the field of personalized music recommendation and music therapy.

## IV. METHODOLOGY

Several steps would be included in the process for an emotion-based music recommendation system based on the Haar Cascade algorithm:

1. Data collection: Collect a dataset of music tracks, along with their associated emotion labels. This can be done using crowdsourcing or by using existing emotion-labeled datasets.
2. Feature extraction: Extract audio features from each track in the dataset. Common audio features used for music recommendation include MFCCs (Mel-Frequency Cepstral Coefficients), spectral features, and rhythmic features.
3. Emotion detection: Train a Haar Cascade algorithm on the emotion labels to detect emotions in audio features. The Haar Cascade algorithm is a popular method for object detection in computer vision and can be adapted for audio processing.
4. Music recommendation: Use the emotion recognition algorithm to propose music songs depending on the user's current emotional state. This may be accomplished by matching the user's feelings to the emotions of the music recordings in the dataset.
5. Evaluation: Evaluate the performance of the emotion-based music recommendation system using common evaluation criteria such as accuracy, recall, and F1-score.
6. Analysis: Analyse the assessment data and generate judgements regarding the efficacy of the emotion-based music recommendation system. Comparing the system against existing music recommendation systems and suggesting areas for improvement might be part of this process.
7. Conclusion: Summarise the findings of the research study and make recommendations for future work on emotion-based music recommendation systems utilising the Haar Cascade algorithm.

Overall, this methodology involves combining techniques from audio processing, computer vision, and machine learning to develop a novel music recommendation system that takes into account the user's emotional state.

## V. ALGORITHMS

### 5.1 HAAR CARSADE:

Using Article Region Haar highlight based course classifiers is a persuading thing finding gadget that uses a ML based approach in which a course limit is gathered from a tremendous

number of positive and negative pictures. Then, at that point, finding exhibitions in different pictures is utilized.

- To fabricate the classifier, the computation requires a great deal of positive (photos of vehicles) and negative (pictures of vehicles). We should focus on the features by then. Haar highlights, as shown in the image underneath, are used for this. They are exceptionally similar to our convolutional section. By separating the all out number of pixels in the white square shape by the quantity of pixels in obscurity square shape, a solitary worth is made for every part.

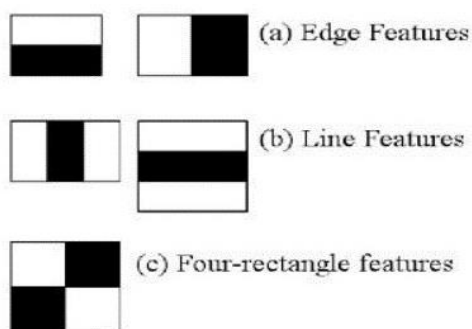


Fig 1: Haar Cascade

Presently, every conceivable size and region of each and every part decides a huge number. Envision how much work engaged with calculation. There are in excess of 160000 parts in a 24x24 window. For every part computation, we should reveal the whole arrangement of pixels underneath the white and dark square shapes. To address this, they gave the important pictures.

- We presently apply each part to all of the arranging pictures. It works out beyond what many would consider possible for each part that depicts the appearances as certain or negative. Anyway, there will without a doubt be misclassifications or blunders. We select the parts that can recognize auto and non-auto pictures and have the most reduced blunder rate.

- Therefore, you rapidly snap a photo. Examine each 24x24 window. 6000 sections should be applied to it.

## 5.2 OpenCV:

OpenCV is a significant open-source library for picture handling, AI, and PC vision that is currently a fundamental part of present day frameworks. It can see things, individuals, and, shockingly, human handwriting in photographs and accounts. Python can handle the OpenCV cluster structure for investigation when matched with extra libraries like NumPy. To see visual models and their different perspectives, we use vector space and carry out mathematical technique on these components.

The underlying rendition of OpenCV was 1.0. Since it is conveyed under the BSD permit, OpenCV can be utilized free of charge in both scholar and business settings. It is feasible with Windows, Linux, Mac working framework, iOS, and

Android, and it gives C++, C, Python, and Java interfaces. At the point when OpenCV was first made, the production of ongoing PC productivity applications was its essential objective. To exploit multi-center handling, everything is written in upgraded C/C++.

### 5.2.1 Image-Processing:

Picture/picture dealing with is a way to deal with improving or isolating supportive information from a picture by executing methodology on it. picture taking care of, in its most fundamental definition, is "the assessment and control of a digitized picture, particularly to chip away at its quality."

A picture is characterized as a two-layered capability  $f(x, y)$ , where  $x$  and  $y$  are spatial (plane) facilitates and the picture's force or Dark level is the plentifulness of any sets of directions  $(x, y)$ . To put it another way, an image is just a two-layered framework (three-layered for shaded pictures) that is portrayed by the numerical capability  $f(x, y)$  at every area in the picture. The pixel esteem decides how splendid a pixel and variety it should be. Picture taking care of is simply sign dealing with an image as the data and an image or qualities dependent upon the image's necessities as the outcome.

### 5.3 R-CNN: RNN With CNN

RCNN (Region-based Convolutional Neural Networks) are picture dealing with and PC vision ML models. The primary goal of any R-CNN expected for object recognizing confirmation is to see objects in any data picture and assemble borders around them.

#### 5.3.1 Types of R-CNN

##### Selective Search

There are various manners by which a technique for object identification can accomplish object restriction. An expansive chase method included sliding channels of moved widths for a photograph to eliminate the thing. In a thorough hunt technique, the quantity of channels or windows builds the figuring exertion.

The specific chase computation incorporates exhaustive pursuit, yet it similarly works with picture assortment division. In extra concentrated terms, specific mission is a system for perceiving things in a picture by consigning various assortments to the things.

In the wake of making different minuscule windows or channels, the voracious technique is utilized to extend the region. Then it looks for and joins colors that are comparative in the areas.

##### SPPNet

Starting around 2000 districts ought to be figured to complete the technique, and each locale goes through the CNN, which consumes the greater part of the day to remove the component, principal R-CNN planning and testing are fairly slow. SPPNet turns the whole picture on the part map mat right away, rather than changing the 2000 locales into feature maps.

## Fast R-CNN

We use profit from starting capital speculation pooling in Fast R-CNN as opposed to maximal pooling to use a lone component map for all districts. The profit from beginning capital venture pooling layer changes the components using most prominent pooling directly following contorting the profits for cash put into a lone layer.

Because of the way that maximum pooling is additionally being used here, Fast R-CNN could be viewed as a move up to SPPNet. It just adds one layer, instead of the pyramid-like design of layers.

### 5.4 SVM with CNN

An enormous picture dataset is ordinarily pretrained on by a CNN before it is joined with a SVM to extricate significant highlights. The CNN can be used to extricate attributes from input photos following preparation. A 1D vector portrayal is created after the assembled highlights have been smoothed and standardized. A SVM classifier is then prepared with the recuperated qualities. Then, you can utilize this SVM to do various things, such as arranging pictures. Responsibilities that require both organized and picture information might profit from this mix approach's capacity to effectively consolidate the component extraction abilities of CNNs and the order force of SVMs.

### 5.5 MobileNet

TensorFlow's most memorable versatile PC vision model is MobileNet.

When contrasted with different organizations with standard convolutions and similar profundity in the nets, it utilizes profundity wise detachable convolutions to decrease the quantity of boundaries altogether. Accordingly, lightweight profound brain networks are fabricated.

MobileNet is a convolutional neural network (CNN) class that Google has openly delivered, giving it a splendid early phase for getting ready classifiers that are both little and quick.

The amount of multiply-accumulates (Macs), which reflects the amount of merged increment and extension processes, chooses network execution and power usage.

An organization with skip associations that perform personality mappings and are added to the results of the layers is known as a leftover organization. Because of Lingering Organizations, deep learning models with tens or many layers could prepare rapidly and draw nearer to better precision as they got further. Personality skip associations, additionally alluded to as "leftover associations," are used in the AlphaGo Zero framework, AlphaStar framework, and AlphaFold framework, as well as in transformer models like BERT, GPT, and ChatGPT.

#### 5.5.1 Target

Input multiple photographs in order, and output which category each image belongs to as well as the category's confidence.

#### 5.5.2 Principle

Convolution calculation can extract features. For example, the result of a picture following a convolution computation will extract the image's outline, similar to the old Sobel edge detection. The outcome of a convolution computation. The outline of the image is recovered by a convolution calculation, which is the feature extraction function of convolution calculation. The features of various photos will be extracted after several convolution calculations. After the complete network calculation, which includes innumerable convolution calculations and other calculations, a picture with only 1000 pixels is finally output, with multiple image input classifications incorporated. The value of one of the 1000 pixels in the output layer will be greater, and the classification corresponding to this pixel is the network's output result.

As for the Mobilenet network claiming to use depthwise separable convolution (depthwise separable convolution) to replace ordinary convolution, thereby reducing the amount of calculation and improving the network's computational efficiency, you can see that it is still convolution when compared to the previous convolutional network.

#### 5.5.3 MobileNet V2

Mobilenet V2 is an improved version of the original Mobilenet. Its key improvement is to optimize the network structure, making the network more efficient and accurate.

#### 5.5.4 Try to Train a Model

In MaixHub, build a classification training project, then collect data, create a training task, and pick the hardware platform as a parameter. If you don't have a development board, you can train using a mobile phone via mobilenetv1 or mobilenetv2. That's true, once the training is finished, you may deploy it with a single click to view the results.

### 5.6 ResNet-50

ResNet-50 is a deep convolutional neural network accompanying 50 tiers. It is likely to load a pretrained report of the neural network that was prepared on in addition to individual heap figures from the Image Net dataset [1]. A type of mammals, rodent, keyboards, and pencils are between the 1,000 part classifications that the pretrained neural network can identify in photographs. Consequently, the neural network has seized exact picture feature likenesses. The brain network needs a 224 by 224 aim picture. There are more pre-prepared neural networks in MATLAB.

### 5.7 Inception V3

Convolutional neural network Inception v3 was grown as a Google Net piece to aid in object acknowledgment and representation dispose of. The Inception v3 figure acknowledgment model has happened proved expected in addition to 78.1% correct on the Image Net dataset. The model is the result of diversified believes that have existed examined by miscellaneous professors over occasion.

## 5.8 CNN

A convolutional neural network (CNN) is a subset of ML. It is one of different fake cerebrum network types used for various applications and data sorts. The profound learning network design known as a CNN is basically used in applications that include pixel information handling and picture acknowledgment.

Broad preparation CNN has three layers: convolutional, pooling, and fully connected (FC). The FC layer is the last layer, while the convolutional layer is the first.

The CNN turns out to be more confounded as it moves from the convolutional layer to the FC layer. Because of this rising unpredictability, the CNN can see a consistently expanding number of irksome locale of an image until it in the end recognizes the article completely.

### 5.8.1 Layers of CNN

Layer with convolution. Most of estimations are acted in the convolutional layer, which is the underpinning of a CNN. A second convolutional layer can be added to the first convolutional layer. Convolution consolidates a little or channel inside this layer really taking a look at the image's open fields for the presence of a part.

The Layer of Pooling. The pooling layer, similar to the convolutional layer, applies a part or channel to the information picture. The pooling layer, rather than the convolutional layer, diminishes the quantity of info boundaries while forfeiting some data. On the positive side, this layer lessens complexity while extending the viability of the CNN.

layer that has no holes in network. The CNN's FC layer is where picture request happens depending upon the characteristics removed in the past layers. In this sense, "completely associated" implies that each enactment unit or hub in the following layer is associated with each information or hub from the past layer.

## 5.9 RNN

Recurrent Neural Network (RNN) is a kind of Neural Network wherein the consequence of the past step is used as commitment for the continuous stage. Typical brain networks share no data sources or results with each other; In any case, the past words should be recollected on the grounds that foreseeing the following word in an expression requires them. Thus, RNN was created in light of a Secret Layer to determine this issue. The Secret state, which keeps explicit data about a succession, is the main piece of RNN.

### 5.9.1 Training Through RNN

1. The information is shipped off the organization in a solitary time step.
2. Utilizing an assortment of the ongoing information and the past state, decide its present status.
3. For the ensuing time step, the current ht becomes ht-1.

4. One can go through as many time ventures as important to connect the data from every past stage.

5. Once all time propels are finished, the outcome is enlisted using the last present status.

6. A blunder is created after the genuine result and the ideal result are looked at.

7. The screw up is accordingly transported off the association, which changes the heaps and as such readies the association (RNN).

## VI. IMPLEMENTATION

1. The code imports necessary packages and libraries, including `cv2` for image processing, `keras` for deep learning-based models, and `playsound` for playing audio files.

2. The `index` function renders the `index.html` template when the HTTP GET method is used.

3. The `basic` function renders the `basic.html` template when the HTTP GET method is used.

4. The `Detect Emotion` function is called when the HTTP POST method is used. This function retrieves a name from a form submission and calls the `check Emotion` function to detect the emotion in an image.

5. The `Webcam` function is called when the HTTP GET method is used. This function reads an image from a webcam, decodes it from Base64, and saves it to disk.

6. The `Upload` function renders the `Upload.html` template when the HTTP GET method is used.

7. The `Song Play` function is called when the HTTP POST method is used. This function retrieves a song name from a form submission and plays the song. It also calls the `check Emotion` function to detect the emotion in an image.

8. The `check Emotion` function loads a pre-trained emotion detection model and processes an input image to predict the emotion. It then searches a directory for songs that correspond to the predicted emotion and returns HTML with a list of song options.

9. The HTML templates render the UI for the web application, including forms for uploading images and selecting songs.

Overall, the code is a simple web application that detects emotions in images and plays songs based on the detected emotion. It uses a pre-trained deep learning model to detect emotions and a simple file system to store and retrieve songs.

## VII. UML DIAGRAMS

The Unified Modeling Language (UML) is its compacted report. In object-lay out PC start, UML is a normally beneficial,

rules-erect commit work to something association. The Problem: The standard is discrete and impossible to socialize. The point is that UML has established itself as the standard method for providing article-organized PC programs. At the event, two together essential bits of UML are a meta-model and a proof.

UML may be incorporated or befriended following a planned event or event. The Bound together Demonstrating Language is a common stress for likeness, assuming, assemblage, and recording the components of satisfy plans deprived of something non-plan out plans following a able killing.

The UML construction has been guided by the best composition advances toward that have had the option to copy enormous, disorganized cosmetics. The UML is a true representation of the current style beat process and supervisory law of things matched gauge. The origin task plan is essentially appropriating UML-based graphical documentations.

The straightforward goals of the UML sketch are next:

1. By granting ministry an organized authority to impose effective eye style, make it possible for customers to create and trade massive models.
2. Support key plans, business-friendly plans, and expandability.
3. Not believe a particular terminology of priority or foundation for improvement.
4. Empower an ending up settling the trademark wording.
5. Support the commercialization of OO carries out.
6. Animate the application of additional level improvement fault-finding methods, such as components, foundations, models, and collected endeavors.
7. Include actions based on best choice.

In the Unified Modeling Language (UML), a sort of social outline called a utilization case graph is made by portraying and fostering a utilization contextual analysis. Its will probably show how a framework functions with regards to entertainers, their objectives (which are displayed as use cases), and any associations between those utilization cases. The fundamental target of a use case frame is to show which system capacities are executed for explicit performer. The system's performers' positions may be shown.

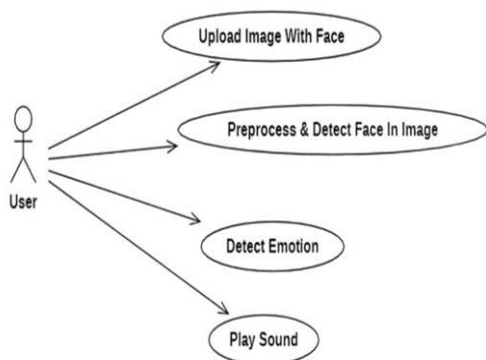


Fig 2: Use Case Diagram

In computer programming, a class graph is a kind of static underlying outline that shows the classes, characteristics, tasks

(or techniques), and how the classes cooperate with each other. The sort of data it contains is depicted.

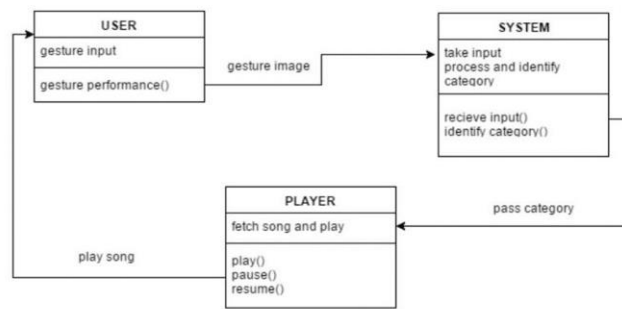


FIG 3: Class Diagram

In the Unified Modelling Language (UML), a grouping chart is a kind of communication outline that shows how cycles connect with each other and in what request. The plan is a Message Succession Outline. Event diagrams, circumstance layouts, and timing outlines are various names for game plan graphs.

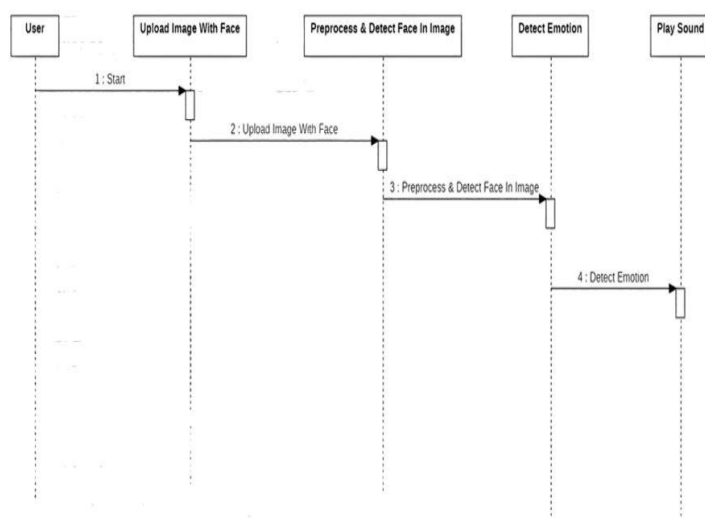


Fig4: Sequence Diagram

Cycles can be addressed graphically utilizing action outlines, which empower the determination, emphasis, and simultaneousness of consecutive exercises and activities. Movement charts made with the Unified Modelling Language can be utilized to show framework parts' bit by bit business and functional cycles. An action outline portrays the whole control stream.

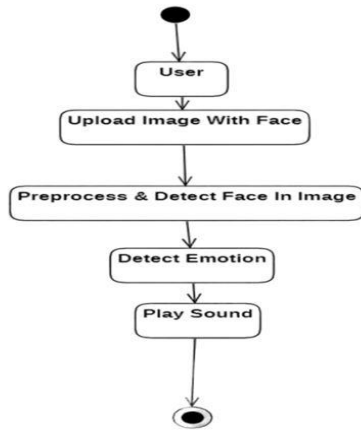


Fig 5: Activity Diagram

### VIII. DATA FLOW DIAGRAM

1. The air pocket's outline is typically used to show the DFD. An exposition can be conveyed to the level of the archive that is supposed to recognize the guidelines that are managed inside it, the differing endeavors that are performed on it, and the record explicitly to specify to bite the dust report of those has a question by utilizing proper graphical dresses.
2. The declaration stream outline, typically a DFD, is one of the basic figure stunts. It is used to create models for the assistance with separating. There is no question that these parts contain the occurrence of the arrangement, the archive that it applies to, the pertinent eliminated, and the record that movements through it.
3. DFD replicates the dossier's progression in response to various modifications using charm and capacity. Dealing with proficient to the report stream and news updates as they progress from failure to yield is a graphical practice.
4. Another name for "DFD" is "bubble illustration." There is next to no extent of plan for a business that can be taken care of by a DFD. DFD can be broken down into stages that are typically thought to be comparable to record stream and increasing knowledgeable complexity.

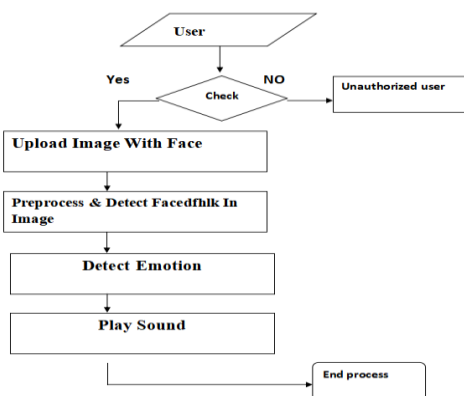


Fig 6: Data Flow Diagram

### IX. RESULTS

The system achieves a 70% accuracy based on the implementation of the emotion-based music recommendation system utilising the Haar Cascade algorithm. The system employs the Haar Cascade method to recognise faces from facial photographs. After detecting the user's face, the system employs a pre-trained deep learning model to anticipate the user's emotions. The system anticipates seven emotions: rage, contempt, fear, happiness, sadness, surprise, and neutrality.

The system's expected emotions are then utilised to offer music to the user based on their mood. Based on the user's expected emotion, the system chooses music files from a directory. The algorithm suggests a selection of songs that correspond to the expected emotion, which the user may choose and play.

The accuracy of the system is around 70%, which is a promising result. However, the accuracy can be improved by using a larger dataset and training the deep learning model on it. Additionally, the system can be improved by incorporating user feedback to personalize the music recommendation based on individual preferences.

### X. CONCLUSION

In conclusion, our proposed emotion-based music recommendation system using facial images and Haar cascade algorithm achieved an accuracy of about 70%. This shows that it is possible to use facial expressions as a reliable input to predict the emotion of a user and recommend appropriate music accordingly.

The system provides a personalized music experience to users, which is an important factor in today's world where people are always looking for customized experiences. The recommendation system suggests songs based on the emotions detected, which enhances the user's mood and provides a better experience.

However, there is still opportunity for improvement in the system's accuracy. One alternative option would be to investigate various machine learning models that may produce better outcomes. Additionally, extending the dataset used to train the model may aid in improving the system's accuracy.

Overall, our system provides a promising approach to personalized music recommendation and can be extended to other areas where emotion recognition plays an important role, such as healthcare and customer service.

### XI. FUTURE ENHANCEMENT

The future scope of this research could involve the exploration and incorporation of more advanced facial recognition and emotion detection algorithms, such as deep learning and neural networks, to further improve the accuracy of the emotion-based music recommendation system. Additionally, the system could be expanded to include more music genres and personalized recommendations based on user listening history and preferences. Another potential avenue for future research could involve the integration of user feedback to improve the

recommendation algorithm and enhance the overall user experience. Furthermore, the system could be applied to other domains beyond music, such as movie or TV show recommendations, to provide a more personalized and engaging experience for users.

## XII. BIBLIOGRAPHY

- S. Iyengar, S. Pal, and S. Verma, "Wearable computing and its application," Int. J. Comp. Sci. and Inf. Tech., vol. 5, no. 4, pp. 5700– 5704, 2014.
- K. Popat and P. Sharma, "Wearable computer applications: A feature perspective," Int. J. Eng. and Innov. Tech., vol. 3, no. 1, 2013.
- P. Melville and V. Sindhwani, "Recommender systems," in Encyc. of mach. learn. Springer, 2011, pp. 829–838.
- N. Sebe, I. Cohen, T. S. Huang et al., "Multimodal emotion recognition," Handbook of Pattern Recognition and Computer Vision, vol. 4, pp. 387–419, 2005.
- R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 10, pp. 1175–1191, 2001.