MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE GWALIOR

(A Govt. Aided UGC Autonomous Institute Affiliated to RGPV, Bhopal)

NAAC Accredited with A++ Grade

Project Report

on

"Emotions based music recommendation system"

Submitted By:

Manish Sharma

(0901AD211028)

Gaurav Chaudhary

(0901AD211015)

Faculty Mentor:

Mr. Mir Shahnawaz

CENTRE FOR ARTIFICIAL INTELLIGENCE

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE

GWALIOR - 474005 (MP) est. 1957

JULY-DEC. 2023

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE GWALIOR

(A Govt. Aided UGC Autonomous Institute Affiliated to RGPV, Bhopal)

NAAC Accredited with A++ Grade

CERTIFICATE

This is certified that **Manish Sharma (0901AD211028)** and **Gaurav Chaudhary (0901AD211015)** has submitted the project report titled **Emotions based music recommendation system** under the mentorship of **Mr. Mir Shahnawaz** Ahmad, in partial fulfilment of the requirement for the award of degree of Bachelor of Technology in **Artificial intelligence and data science** from Madhav Institute of Technology and Science, Gwalior.

Mr. Mir Shahnawaz	**Dr. R. R. Singh**
Faculty Mentor	Coordinator
Centre for Artificial Intelligence	Centre for Artificial Intelligence

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE GWALIOR

(A Govt. Aided UGC Autonomous Institute Affiliated to RGPV, Bhopal)

NAAC Accredited with A++ Grade

DECLARATION

I hereby declare that the work being presented in this project report, for the partial fulfilment of requirement for the award of the degree of Bachelor of Technology in **Artificial intelligence and data science** at Madhav Institute of Technology & Science, Gwalior is an authenticated and original record of my work under the mentorship of **Mr. Mir Shahnawaz** Centre for Artificial Intelligence.

I declare that I have not submitted the matter embodied in this report for the award of any degree or diploma anywhere else.

Manish Sharma (0901AD211028)
Gaurav Chaudhary (0901AD211015)
3rd year, Artificial intelligence and data science (AI&DS)
Centre for Artificial Intelligence

MADHAV INSTITUTE OF TECHNOLOGY & SCIENCE GWALIOR

(A Govt. Aided UGC Autonomous Institute Affiliated to RGPV, Bhopal)

NAAC Accredited with A++ Grade

ACKNOWLEDGEMENT

The full semester project has proved to be pivotal to my career. I am thankful to my institute, **Madhav institute of Technology and Science** to allow me to continue my disciplinary/interdisciplinary project as a curriculum requirement, under the provisions of the Flexible Curriculum Scheme (based on the AICTE Model Curriculum 2018), approved by the Academic Council of the institute. I extend my gratitude to the Director of the institute, **Dr. R. K. Pandit** and Dean Academics, **Dr. Manjaree Pandit** for this.

I would sincerely like to thank my department, **Centre for Artificial Intelligence,** for allowing me to explore this project. I humbly thank **Dr. R. R. Singh**, Coordinator, Centre for Artificial Intelligence, for his continued support during the course of this engagement, which eased the process and formalities involved.

I am sincerely thankful to my faculty mentors. I am grateful to the guidance of **Mr. Mir Shahnawaz**, <Designation>, Centre for Artificial Intelligence , for his continued support and guidance throughout the project. I am also very thankful to the faculty and staff of the department.

Manish Sharma (0901AD211028)
Gaurav Chaudhary (0901AD211015)
3rd year, Artificial intelligence and data science (AI&DS)
Centre for Artificial Intelligence

ABSTRACT

A user's emotion or mood can be detected by his/her facial expressions. These expressions can be derived from the live feed via the system's camera. A lot of research is being conducted in the field of Computer Vision and Machine Learning (ML), where machines are trained to identify various human emotions or moods. Machine Learning provides various techniques through which human emotions can be detected. One such technique is to use MobileNet model with Keras, which generates a small size trained model and makes Android-ML integration easier.

Music is a great connector. It unites us across markets, ages, backgrounds, languages, preferences, political leanings and income levels. Music players and other streaming apps have a high demand as these apps can be used anytime, anywhere and can be combined with daily activities, travelling, sports, etc. With the rapid development of mobile networks and digital multimedia technologies, digital music has become the mainstream consumer content sought by many young people. People often use music as a means of mood regulation, specifically to change a bad mood, increase energy level or reduce tension. Also, listening to the right kind of music at the right time may improve mental health. Thus, human emotions have a strong relationship with music. In our proposed system, a mood-based music player is created which performs real time mood detection and suggests songs as per detected mood. This becomes an additional feature to the traditional music player apps that come pre-installed in our mobile phones. An important benefit of incorporating mood detection is customer satisfaction. The objective of this system is to analyse the user’s image, predict the expression of the user and suggest songs suitable to the detected mood.

Keyword: Face Recognition, Image Processing, Computer Vision, Emotion Detection, Music, Mood detection

सारः

एक उपयोगकर्ता की भावना या मूड को उसके चेहरे के अभिव्यक्तियों से पहचाना जा सकता है। इन अभिव्यक्तियों को सिस्टम के कैमरा के माध्यम से लाइव फीड से प्राप्त किया जा सकता है। कंप्यूटर विजन और मशीन लर्निंग के क्षेत्र में बहुत से अनुसंधान हो रहे हैं, जहां मशीनों को विभिन्न मानव भावनाओं या मूड की पहचान करने के लिए प्रशिक्षित किया जा रहा है। मशीन लर्निंग विभिन्न तकनीकें प्रदान करता है जिनके माध्यम से मानव भावनाएं पहचानी जा सकती हैं। एक ऐसी तकनीक में, मोबाइलनेट मॉडल का उपयोग करना एक विकल्प है जो केरस के साथ किया जा सकता है, जो एक छोटे आकार के प्रशिक्षित मॉडल उत्पन्न करता है और एंड्रॉइड-एमएल एकीकरण को सरल बनाता है।

संगीत एक महान कनेक्टर है। यह हमें बाजारों, आयु, पृष्ठभूमियों, भाषाओं, पसंदों, राजनीतिक धाराओं और आय स्तरों के बीच एकजुट करता है। संगीत प्लेयर्स और अन्य स्ट्रीमिंग ऐप्स की उच्च मांग है क्योंकि इन ऐप्स का कभी भी, कहीं भी उपयोग किया जा सकता है और इन्हें दैनिक गतिविधियों, यात्रा, खेल आदि के साथ संयुक्त किया जा सकता है। मोबाइल नेटवर्क्स और डिजिटल मल्टीमीडिया तकनीकों के शीघ्र विकास के साथ, डिजिटल संगीत ने कई युवा लोगों द्वारा खोजे जाने वाले प्रमुख उपभोक्ता सामग्री बन गया है। लोग अक्सर संगीत का उपयोग भावना नियंत्रण के रूप में करते हैं, विशेषकर एक बुरे मूड को बदलने, ऊर्जा स्तर बढ़ाने या तनाव को कम करने के लिए। यहां तक कि सही समय पर सही प्रकार का संगीत सुनने से मानसिक स्वास्थ्य में सुधार हो सकती है। इस प्रस्तावित सिस्टम में, एक मूड-आधारित संगीत प्लेयर बनाया गया है जो वास्तविक समय में मूड पहचान करता है और पहचाने गए मूड के अनुसार गाने सुझाए जाते हैं। यह हमारे मोबाइल फोन में पूर्व-स्थापित संगीत प्लेयर ऐप्स के लिए एक अतिरिक्त सुविधा बन जाता है। मूड पहचान को शामिल करने का एक महत्वपूर्ण लाभ ग्राहक संतुष्टि है।

TABLE OF CONTENTS

Chapter 1: Introduction

1.1 Description:

Human emotions can be broadly classified as: fear, disgust, anger, surprise, sad, happy and neutral. A large number of other emotions such as cheerful (which is a variation of happy) and contempt (which is a variation of disgust) can be categorized under this umbrella of emotions. These emotions are very subtle. Facial muscle contortions are very minimal, and detecting these differences can be very challenging as even a small difference results in different expressions. Also, expressions of different or even the same people might vary for the same emotion, as emotions are hugely context dependent. While the focus can on only those areas of the face which display a maximum of emotions like around the mouth and eyes, how these gestures are extracted and categorized is still an important question. Neural networks and machine learning have been used for these tasks and have obtained good results. Machine learning algorithms have proven to be very useful in pattern recognition and classification, and hence can be used for mood detection as well.

With the development of digital music technology, the development of a personalized music recommendation system which recommends music for users is essential. It is a big challenge to provide recommendations from the large data available on the internet. E-commerce giants like Amazon, EBay provide personalized recommendations to users based on their taste and history while companies like Spotify, Pandora use Machine Learning and Deep Learning techniques for providing appropriate recommendations. There has been some work done on personalized music recommendation to recommend songs based on the user's preference. There exist two major approaches for the personalized music recommendation. One is the content based filtering approach which analyses the content of music that users liked in the past and recommends the music with relevant content. The main drawback of this approach is that the model can only make recommendations based on existing interests of the user. In other words, the model has limited ability to expand on the users' existing interests. The other approach is the collaborative filtering approach which recommends music that a peer group of similar preference liked. Both recommendation approaches are based on the user's preferences observed from the listening behaviour. The major drawback of this approach is the popularity bias problem: popular (i.e., frequently rated) items get a lot of exposure while less popular ones are under-represented in the recommendations. Generally, a hybrid approach is implemented in which both content and collaborative techniques are combined to extract maximum accuracy and to overcome drawbacks of both types. [1]

In this work, the aim is to create a music recommendation system/music player which will detect the user's face, identify the current mood and then recommend a playlist based on the detected mood.

1.2 Literature Review:

In a particular system [8], Anaconda and Python 3.5 softwares were used to test the functionality and Viola Jones and haar cascade algorithms were used for face detection. Similarly, KDEF (Karolinska Directed Emotional Faces) dataset and VGG (Visual Geometry Group) 16 were used with CNN (Convolution Neural Network) model which was designed with an accuracy of 88%, for face recognition and classification that validated the performance measures. However, the results proved that the network architecture designed had better advancements than existing algorithms. Another system [9] used Python 2.7, Open Source Computer Vision Library (OpenCV) & CK (Cohn Kanade) and CK+ (Extended Cohn-Kanade) database which gave approximately 83% accuracy. Certain researchers have described the Extended Cohn-Kanade (CK+) database for those wanting to prototype and benchmark systems for automatic facial expression detection. The popularity and ease of access for the original Cohn-Kanade dataset this is seen as a very valuable addition to the already existing corpora. It was also stated that for a fully automatic system to be robust for all expressions in a myriad of realistic scenarios, more data is required. For this to occur very large reliably coded datasets across a wide array of visual variabilities are required (at least 5 to 10k examples for each action) which would require a collaborative research effort from various institutions.

There has also been research done on the Music Recommendation System. According to one such research [11], a preliminary approach to Hindi music mood classification has been described, that exploits simple features extracted from the audio. MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy gave an average accuracy of 51.56% using the 10-fold cross validation. In addition to this, there is an article [10] that states that the current music recommendation research results from the perspective ofmusic resources description. It is suggested that there is a lack of systematic research on user behaviour and needs, low level of feature extraction, and a single evaluation index in current research. Situation was identified to be an important factor in the music personalized recommendation system. Finally, it was concluded that when the weights given to all the contextual factors were the same, greatly reduced the accuracy of the recommendation results.

Another research [12] states that their hybrid recommendation system approach concept will work once their model is trained enough to recognize the labels. The mechanism for the automatic management of the user preferences in the personalized music recommendation service automatically extracts the user preference data from the user’s brain waves and audio features from music. In their study, a very short feature vector, obtained from low dimensional projection and already developed audio features, is used for music genre classification problems. A distance metric learning algorithm was applied in order to reduce the dimensionality of the feature vector with a little performance degradation. Proposed user’s preference classifier achieved an overall accuracy of 81.07% in the binary preference classification for the KETI AFA2000 music corpus. The user satisfaction was recognizable when brainwaves were used.

Chapter 2: Components and Theory

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.

2.1 Mood Detection Module

This Module is divided into two parts:

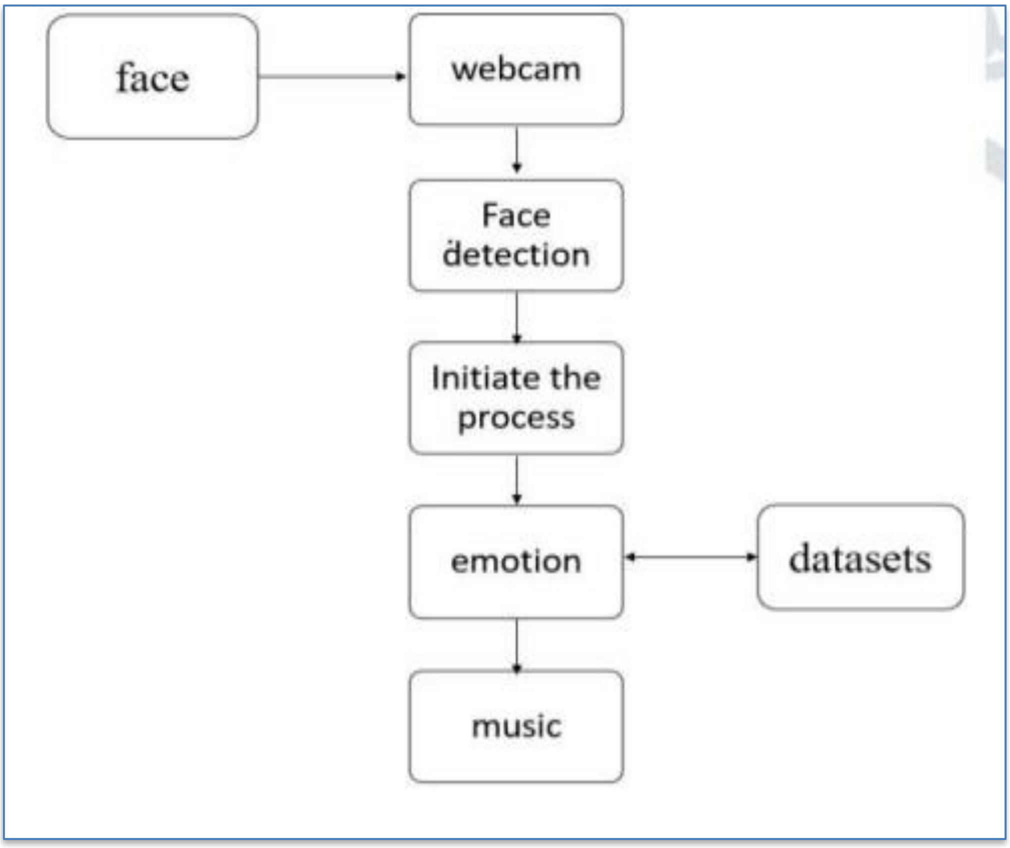
2.1.1 Face Expression Capturing

To capture images, use webcam or any other physiological devices. For that purpose, the python computer vision library is used. This makes it easier to integrate it with other libraries which can also use NumPy, and it is mainly used as a real time computer vision. In the initial process when execution starts it starts to access the camera stream and captures about ten images for further process and emotion detection. So, the initial phase of this work is to capture the images of human face by using Fisher Face Algorithm. Fisher Face Algorithm is used for reducing the face space dimensions using the principal component analysis (PCA) method and then it applies fishers linear discriminant (FDL) or the LDA method to obtain the feature of the image characteristics. This algorithm maximizes the separation between classes in the training process and process for image recognition is done in fisher face. This algorithm use minimum Euclidean to classify the expression that implies the emotion of the user.

2.1.2 Face Emotion Detection

The face recognition is considered as one of the best ways to determine a person’s mood. Fisher face with open CV mainly it mainly emphasis onthe class specific transformation metric by training the model that the value evaluated from the process can help us to deduce the mood of the user. Each emotion is compared with tens of stored images and scale gives the exact emotion. Haarcascade Algorithm is a machine learning algorithm to categorize objects in a captured image. It is mainly used for object detection. Objects here are nose, eyes, ears, lips in face. Haar cascade which is designed by open cv is to detect the frontal face. It also has the capacity to detect the features from the source. It works by training the negative images over the positive images which are superimposed by it. Positive images contain the images only which we want our classifier to categorize, Negative Images contain the Images of everything else, which do not contain the object we want to detect.The cascade classifier has different stages of collection which resembles from weak learners. These weak classifiers are the simplest form classifiers that have a name called boosting. If the label ranges in positive state, then it goes to the next stage showing the result. These have a positive side and a negative side where they identify the images according to the labels. These have a set of positive images over negative images on various stages. As images with higher resolution has greater quantity are preferred as better-quality results.

Flowchart:



2.2 Music Recommendation Module

The input image that is acquired is from the web camera and is used to capture real-time images. It is very hard to define all the emotions and by limited options it can help the compilation time and the outcome is more sophisticated. The emotions are assigned to every song. It compares the values that are present as a threshold in the code. The values will be transferred to perform the web service. The song's will be played from the detected emotion. When the emotion is transferred the respective song and the emotions are numbered are arranged and assigned to every song. However, we can use many kinds of models to recommend because of their accuracy. The fisher face that contains the PCA and LDA algorithms gives the accuracy better than other algorithms and for the sound mechanism win sound (commonly used python library) for basic sound playing for the mechanism obtained are being compared the values that are present as a threshold.

2.3 Hardware and Software Requirements

2.3.1 Hardware:

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. The hardware requirements required for this project are:

• Minimum 4 Gigabyte (GB) RAM (used for processing)

• Webcam (for testing on laptop/desktop)

• Minimum 16 Megapixel (MP) Resolution camera (for testing on android device)

• 30 MB Memory space (approximate value)

2.3.2 Software Requirements

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

• Python 3.6

• OpenCV 3.1

• PyCharm IDE

• Android Studio

Chapter 3: Results

3.1 Happy Emotion

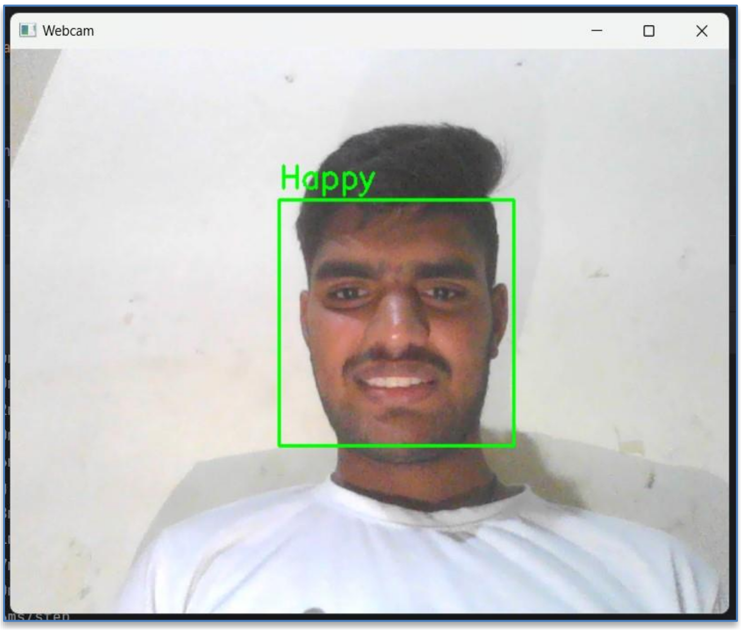


Fig: - Detecting Face and Emotion

```
Detected Emotions: ['Happy', 'Happy', 'Happy', 'Happy', '
Last Detected Emotion : Happy
Video URL: https://www.youtube.com/watch?v=SP07bBQqCDQ
```

Fig: - Printing all Emotion and Capturing Last Emotion

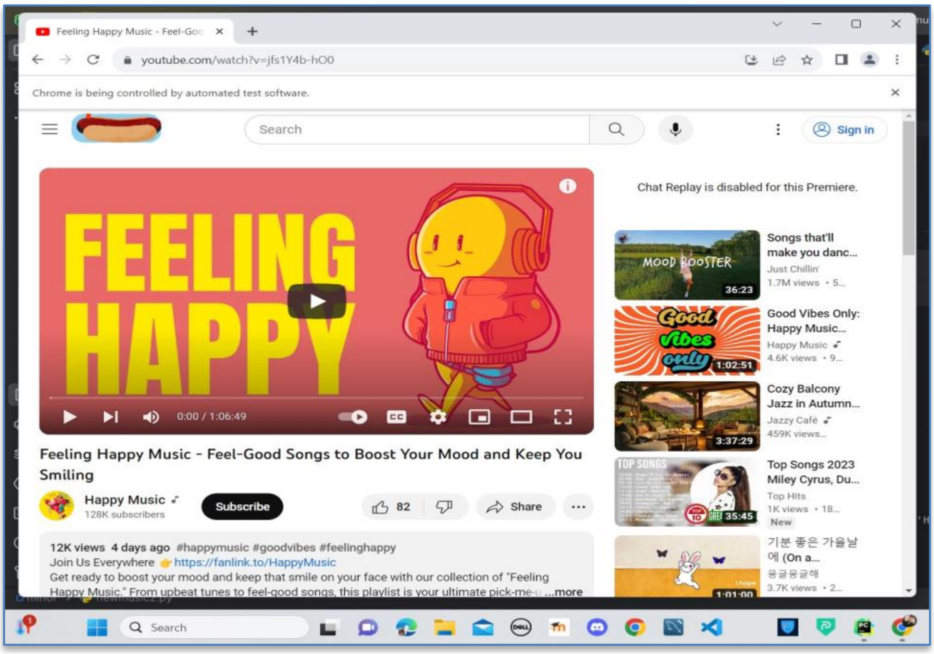


Fig: - Recommend Music Based on Last Emotion

3.2 Neutral Emotions

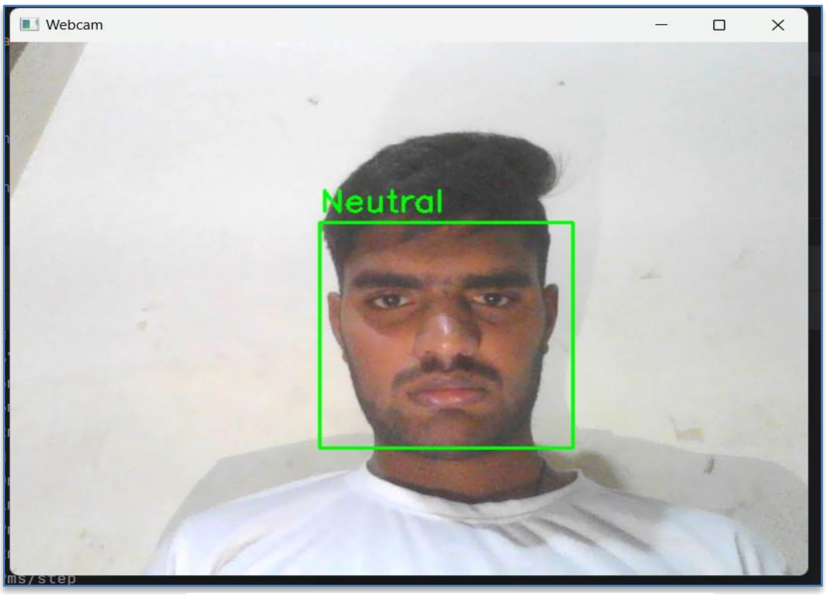


Fig: - Detecting Face and Emotion

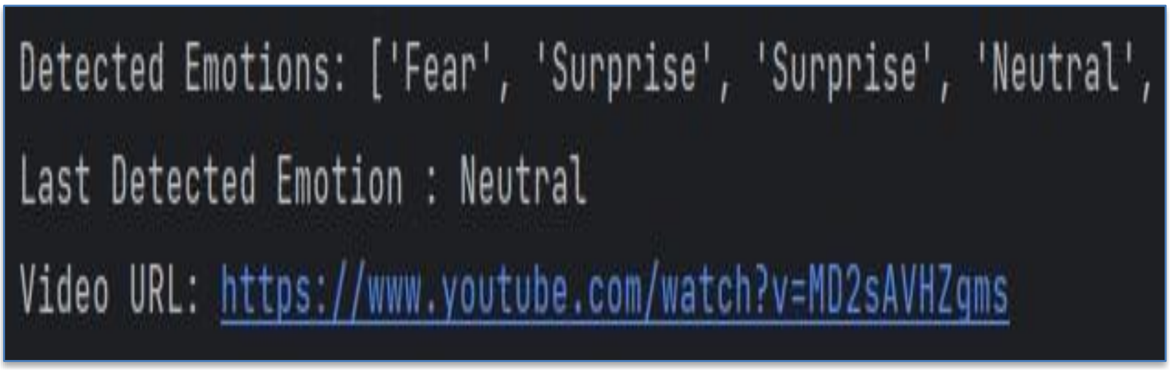


Fig: - Printing all Emotion and Capturing Last Emotion

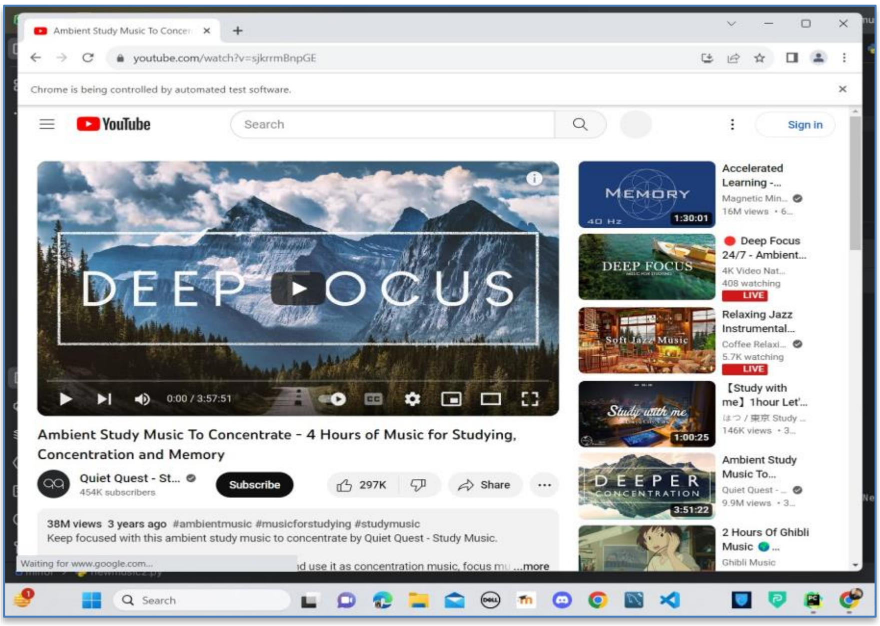


Fig: - Recommend Music Based on Last Emotion

3.2 Angry Emotions

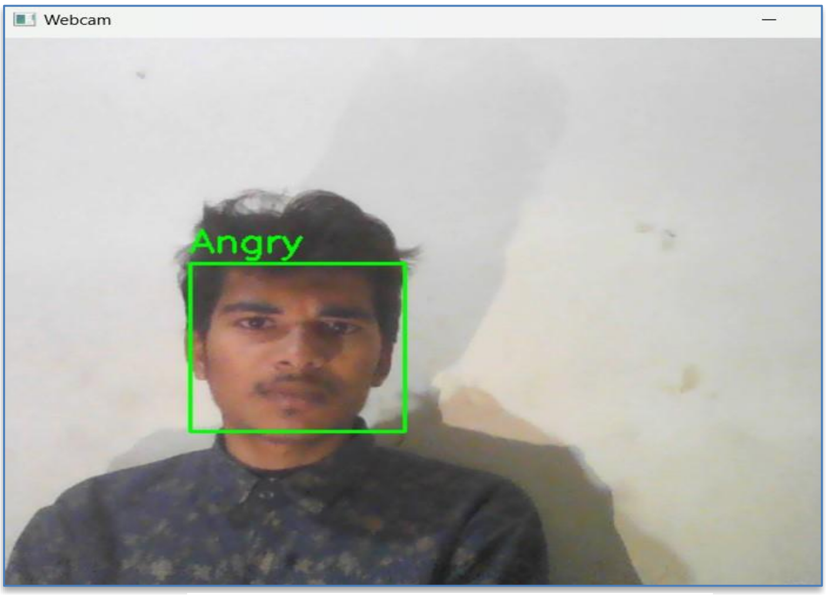


Fig: - Detecting Face and Emotion

Fig: - Printing all Emotion and Capturing Last Emotion

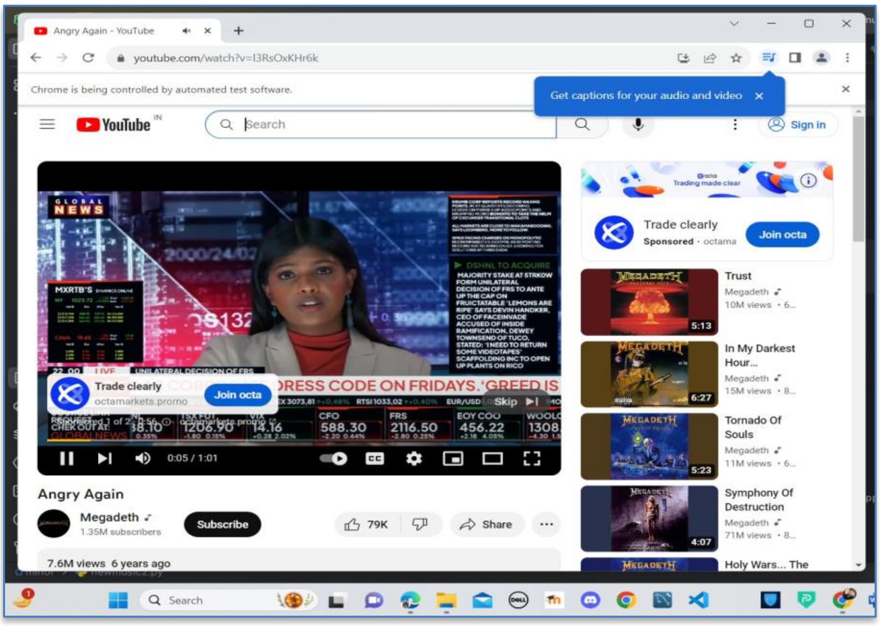


Fig: - Recommend Music Based on Last Emotion

Chapter 4: - Final Analysis and Design:

4.1 Applications: -

an emotions-based music recommendation system has a wide range of applications, enhancing user experience and engagement in various contexts. Here are some applications:

4.1.1: - Personalized Music Streaming Services:

Mood Playlists: Recommend music playlists based on the user's current mood or emotions, creating a personalized and immersive listening experience.

Activity-specific Playlists: Suggest music tailored to specific activities (e.g., workout, relaxation, studying) based on the user's emotional state.

4.1.2: - Enhanced User Engagement:

Emotional Feedback: Provide users with the ability to provide feedback on recommended songs to refine the system's understanding of their emotional preferences over time.

Dynamic Playlists: Create dynamic playlists that adapt to the user's changing emotions, ensuring a continuously relevant and enjoyable music experience.

4.1.3: - Entertainment and Media Production:

Film and Television: Assist in selecting music for scenes in movies and TV shows that align with the desired emotional tone.

Video Games: Enhance gaming experiences by dynamically adapting the music to the emotional context of the gameplay.

4.1.4: - Learning and Productivity:

Focus and Concentration: Recommend music that enhances focus and concentration based on the user's emotional and cognitive state, supporting productivity.

Learning Enhancement: Tailor music recommendations to facilitate learning by aligning with the user's emotional and cognitive needs during educational activities.

Implementing an emotions-based music recommendation system in these areas can significantly improve user satisfaction, engagement, and the overall impact of music on various aspects of life.

4.2: - Scope

The scope of emotions-based music recommendation is broad and encompasses various fields and applications. The continued development and integration of emotion-based algorithms into music recommendation systems open up opportunities in several areas:

4.2.1: - Music Streaming Services:

Market Growth: Emotions-based recommendations can be a key differentiator for music streaming platforms, attracting and retaining users in a competitive market.

Global Appeal: Catering to diverse emotional preferences can make music platforms more appealing on a global scale, considering the cultural and emotional nuances of different regions.

4.2.2: - Entertainment Industry:

Film and TV Production: Emotion-based music recommendations can play a crucial role in selecting and creating soundtracks that amplify the emotional impact of scenes in movies and TV shows.

Video Games: Dynamic music systems in video games can adapt to the player's emotions, enhancing immersion and engagement.

4.2.3: - Education and Productivity:

Learning Environments: Emotion-based music recommendations can be integrated into educational platforms to create an optimal emotional and cognitive environment for learning.

Workplace Productivity: Tailored music playlists can contribute to a positive and focused work environment, potentially boosting productivity.

4.2.4: - Social Media Integration:

User Engagement: Integrating emotions-based music recommendations into social media platforms can enhance user engagement and provide new ways for individuals to express and share their emotions.

4.2.5: - Artificial Intelligence Research:

Algorithmic Advancements: The ongoing development of emotion recognition algorithms contributes to the broader field of artificial intelligence, with potential applications beyond music recommendation.

4.3: - Conclusion

In conclusion, an emotions-based music recommendation system holds significant promise and potential across various domains. By leveraging the emotional context of users, this technology can revolutionize the way people discover and engage with music, creating a more personalized and enriching experience. The applications span from entertainment and leisure to therapeutic and productivity-enhancing contexts.

Such a system not only meets the diverse emotional needs of users but also contributes to the evolution of industries such as streaming services, retail, media production, and beyond. The dynamic adaptation of music playlists based on user emotions ensures that the auditory experience remains relevant and resonant with the current state of mind, promoting a deeper connection between individuals and their music choices.

Furthermore, the integration of emotions-based music recommendations into various aspects of daily life, from social interactions to commercial spaces, demonstrates the versatility and impact of this technology. It has the potential to influence moods, enhance well-being, and contribute to the overall atmosphere in both digital and physical environments.

As technology continues to advance, refining the accuracy and sensitivity of emotions-based music recommendation systems, we can anticipate even more sophisticated and tailored experiences. However, ethical considerations, such as user privacy and consent, should be carefully addressed to ensure the responsible and respectful implementation of these technologies

4.5: - REFERENCES

[1]. AnaghaS.Dhavalikar and Dr. R. K. Kulkarni, “Face Detection and Facial Expression Recognition System” 2014 Interntional Conference on Electronics and Communication System (ICECS -2014)

[2]. Yong-Hwan Lee, Woori Han and Youngseop Kim, “Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting” 2013 16th International Conference on Network-Based Information Systems.

[3]. ArtoLehtiniemi and Jukka Holm, “Using Animated Mood Pictures in Music Recommendation”, 2012 16th International Conference on Information Visualisation.

[4]. F. Abdat, C. Maaoui and A. Pruski, “Humancomputer interaction using emotion recognition from facial expression”, 2011 UKSim 5th European Symposium on Computer.

[5]. T.-H. Wang and J.-J.J. Lien, “Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation,” J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[6] Renuka R. Londhe, Dr.Vrushshen P. Pawar, “Analysis of Facial Expression and Recognition Based On Statistical Approach”, International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[7] AnukritiDureha “An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions” : IJCA 2014.

[8] Bruce Ferwerda and Markus Schedl “Enhancing Music Recommender Systems with Personality Information and Emotional States”: A Proposal: 2014.

[9] P. Melville and V. Sindhwani, “Recommender systems,” in Encyc. of mach. learn. Springer, 2011, pp. 829–838.

[10] 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.