ISSN: 2836-8495

Research Article

Journal of Current Trends in Computer Science Research

Music Recommendation Based on Current Mood Using Ai & Ml

Dharmendra K Roy[1], Anjali. CH[2]*, G. Kavya Sri[3], B. Tharun[4] and K. Venugopal[5]

[1]*Department of Computer Science and Engineering, Faculty of, Assiut University, Assiut, Egypt*

[2]*Department of Computer Science and Engineering, Hyderabad Institute of Technology and Management, Hyderabad, Telangana*

***Corresponding Author**

Anjali.CH, Department of Computer Science and Engineering, Hyderabad Institute of Technology and Management, Telangana.

Submitted: 2023, Apr 20 ; **Accepted:** 2023, May 22 ; **Published:** 2023, June 29

Citation: Roy, K. D., Anjali, C. H., Sri, G. K., Tharun, B., Venugopal, K. (2023). Music Recommendation Based on Current Mood Using Ai & Ml. *J Curr Trends Comp Sci Res, 2*(2), 153-162.

Abstract

Music always had a special connection with our emotions. It's a way to connect people around the world through pure emotion. Yet, it is extremely difficult to generalize about music and say that everyone will like the same kind of music. The mood-based recommendation is very necessary because it can help people relieve stress and listen to calming music based on their current mood. Its main purpose is to accurately predict the user's mood, and then play songs according to the user's choice and current mood. It uses human-computer interaction (HCI) to recognize human emotions from facial images and extracts facial features from the user's face. Whenever a person wishes to listen to a specific kind of music, he might be ending up hearing some other kind of music that doesn't match his mood. So, the primary objective of our project is to address the major challenges that are faced by users while trying to listen to any music at random. In this paper, two approaches are proposed to overcome the above–mentioned challenges. The first approach is implementing a questionnaire model. The user will be required to respond to a series of questions in this model, and depending on his response, the mood will be determined, and music will be suggested. While the second approach involves designing a model to identify user emotions and then suggest music depending on the mood detected. However, if the user is not satisfied with the emotion captured, he will also have the choice to return to one of the two previously described models and continue the procedure. Our project aims at developing a user interactive and user–friendly model.

Keywords: Convolutional neural networks Emotion detection Image processing Music recommendation

1. Introduction

Music and emotion are inseparably connected, and they can affect each other. At the equivalent time, sure styles of track can substitute a people's emotion. Music counsel fundamentally founded absolutely on temper is essentially needed as it will help individuals in easing strain and focusing on calming track principally founded absolutely on their contemporary day sentiments. The undertaking will probably hold onto a singular feeling by means of looks. A track member is intended to hold onto human feeling with the utilization of the PC's web-advanced Digicam trademark.

Our product captures the individual's photograph after which, the utilization of photograph-handling strategies, extricates capacities from detecting the facial features of the user by webcam. face and attempts to stagger on the feeling that the individual is making an endeavor too explicit. Adding excellent perceptual abilities to PC frameworks could allow them to work together as personal mates with individuals. Scientists attempt to introduce PC frameworks with more noteworthy capacities while heading to allow them to have connections with individuals, acknowledge human presents, talk, tune in, or even bet their sentiments. It utilizes the greatest present-day camcorders and amplifiers to find the individual's developments by means of involving conferred tactile abilities in a non-meddling detecting strategy. The framework can understand what an individual needs, wherein he's looking, or even his real or close-to-home states. Thus, the Feeling of Identification. We utilized the AI idea, which consolidates facial examining and trademark following, to conclude the individual's attitude after which they offer a modified playlist essentially founded absolutely on that. The feeling identification module is utilized to find the inclination communicated through the method of a method for the individual, making it imperative for the product to offer relaxation inside the state of Music principally based absolutely on the individual's attitude. The product is parted into 3 areas: questionnaire, music recommendation, and mood recognition.

Music recommendation is required because of how unique our choices and combinations are in terms of music preferences. When all of this is considered, it is exceedingly difficult to generalize music and state that everyone would enjoy the same type. Music always has a unique link to our emotions. It is a way to connect people around the world through pure feelings. Depending on their current mood, everyone would want to listen to diverse categories of music.

1.1 Goals:

Our goal is to rightly combine the times with the human need for entertainment in the form of music. We first learned about the emotions expressed by the user by once an emotion is detected, the app will play a song based on the emotion detected.

The purpose of the system is to analyze the user's images, predict the user's facial expressions and recommend suitable songs based on the detected emotions.

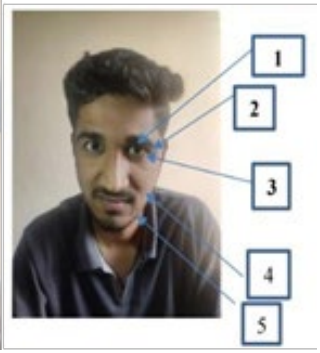

Figure 1 **Figure 2**

Figure 1: The six basic emotions
Figure 2: Distinguishing features of anger

1. Eyebrows pulled down
2. Upper lips pulled up
3. Lower lips pulled up
4. Margins of lips rolled up
5. Lips may be tightened

2. Literature Survey

Anukritin proposed an algorithm that uses the Fisher Face algorithm to list songs in a user's playlist based on their emotions[1]. The designed algorithm focuses on reducing the computation time and therefore the costs related to the use of various hardware. The main idea is to classify emotions into five categories i.e. Joy, sadness, anger, surprise, and fear also provides a highly accurate audio information retrieval method that can extract relevant information from audio signals in less time. Using the Fisher face algorithm, a model can be developed which detects the emotion of the user and classifies the emo,tion and recommends songs.

2.1 Fisher face Algorithm – Multidimensional reduction process that brings together or organizes raw data that has been dimensionality reduced to many other categories. Capture user emotion through web cameras to capture user images. The captured image is enhanced by following the raw data dimensionality reduction process. This data is converted into a binary image format and the Fisher's face and Harcascade methods are used to detect faces. Initial or raw data is extracted from faces and reduced to many other classes. Use the methods described above to sort and organize these classes. Detect emotions by extracting facial features. The main goal of the feature extraction module is to reduce the number of resources required for large datasets. The image features consist of 3 parts

1. Border/Edge
2. Projection corner/point
3. Field Point

Initial or raw data is extracted from faces reduced to this image-processing system using principal component analysis (PCA) to reduce the dimensionality of face space and then applies Fisher's Linear Discriminant method (LDA) or LDA to obtain face features. We use it because it maximizes the separation between classes during training. This algorithm helps to handle the image recognition performed in Fisher's face, while to match faces we use the Euclidean algorithm which helps us classify expressions that suggest the user's emotions.

2.2 ANN Model – Developed by Renuka R Londhe has published an article on face curvature changes and corresponding pixel intensities[2]. The authors used an artificial neural network (ANN) to classify emotions. The author also offers several approaches to playlists. Zheng et al.

2.3 Haarcascade Technique – Model developed using haar cascade technique Nikhil Determining the users' state of mind through facial expressions[3]. People often express their feelings through their facial expressions and gestures but most people express their feelings through their faces. Emotion-based music player reduces time complexity for users, generally,

people have tons of songs in their playlists. Playing random songs does not satisfy the user's mood. The system helps users to automatically play songs according to their mood. The user's image is captured by the webcam and the image is saved. The image is first converted from RGB format to binary format. This process of representing data is called a feature point detection method. This process can also be done using the haar cascade technique provided by open CV. The music player is developed using the Java program. It manages the database and plays songs according to the user's mood.

2.4 Automate the Interaction - Parul Tambe came up with an idea to automate the interaction between the user and the music player, knowing all the preferences, emotions, and activities of the user and giving a selection of songs accordingly, the device records different facial expressions of the user to determine the mood of the user and thus predicting the genre of the music [3].

2.5 Brain-Computer Interface – Chang Liu describes a system using a brain–computer interface, also known as BCI. BCIs use devices to send signals to processing systems [4]. EEG hardware is used to monitor a person's cognitive state. The downside of this scheme is that it constantly requires input from the user's brain to perform the classification. Mid-based algorithms are used to continuously monitor, and process signals received from the user's brain and use the signals to actively monitor and generate the emotions the user is currently experiencing. Swati Vadi reviews EEG – Electroencephalography (EEG) is a form of medical science that records the electrical activity of brain cells and neurons [23]. The electrical activity of neurons in brain cells is recorded. An approximation is made from the recorded neural activity, and a person's emotion is estimated from this analysis. Electroencephalography (EEG) is a form of medical science that records the electrical activity of brain cells and neurons. The electrical activity of neurons in brain cells is recorded.

3. Existing Systems

The recommender systems have become an essential element in the digital world. Users themselves getting confused because of the massive increase in the contents of digital data. The important use cases of the recommendation system include personalizing content, and user-based filtering makes a better product experience for the users.

There are many automatic music recommendation systems that aim to automate the basic operations of applications. However, most of these systems focus directly on recommending the music regardless of user satisfaction. Even though the songs are recommended based on the user's mood, still if the user is not satisfied and wishes to change the mood, it is becoming difficult to find a user-friendly model.

Current music systems are working by AI. The music systems are determining and recommend the music based on three main characteristics the type of lyrical content they are listening to, most frequently listening language, and past listening habits. Apart from these three characteristics, it will be more user-friendly if emotion detection and song recommendation is also added. Then the system will be more interactive for the user.

3.1 Disadvantages of the existing system

1. Recommending songs based on other people's browsing history.
2. Users are unable to find out what they are interested in because of the huge increase in digital content.
3. Recommending songs similar to those already listened to.
4. Unable to filter new songs based on user satisfaction.
5. Lack of personalization and user engagement.

4. Proposed System

This research focuses on designing a model which is more personalized and user-friendly. In order to resolve the above-mentioned challenges, we are going to follow two approaches the first approach is building a 'questionnaire' model and the second approach involves building a 'detect emotion and recommend' model. In our proposed system we are going to give the user two options he can either select the questionnaire option and use the application or select the detect emotion and recommend option and use the application.

Generally, if the user selects the questionnaire option, this questionnaire model consists of a set of questions. The user is required to answer those questions based on the response received the mood of the user will be detected and according to the mood the songs will be recommended. If the user selects the 'detect emotion and recommend' option, then the user's emotion will be captured. Later based on the emotion captured the mood will be detected and music is recommended. If the user is not satisfied with the emotion detected, then he will also have the option to either use any one of the proposed models exactly from where they are without going back.

Our system helps the users in two ways. It can both act as an emotion detection which is so-called a capture and detect model also it can act as a questionnaire model. So, whenever the user register to our website and login to the website. Then he will be able to see our two products which were developed. Based on the user's interests he can go with any of the models. If he is using the detect emotion and recommend model then the user should face the camera, he should show his face to the camera once the web camera got turned on. When the camera is turned on, he will be able to see the emotions, he will also have the flexibility to change the emotion and listen to the songs. However, if he is not interested to face the web camera then he can easily go to the questionnaire model.

4.1 Requirement analysis.

4.1.1 Software: The operating system used will be Windows 9 & above. The programming languages used are Python, HTML5, CSS, and Bootstrap. Tools and frameworks that can be used are Anaconda, Visual Studio community version, Jupiter, Spyder, Pandas, NumPy, Seaborn, Sci-kit learn, Flask, and Tensor flow.

4.1.2 Hardware: Applications that need to store large arrays/objects in memory will require more RAM, whereas applications that need to perform numerous calculations or tasks more quickly will require a faster processor. It requires a web camera. It requires a hard disk of 16GB or more. It requires a processor of pentium4. Ram of size 512 MB will be required.

A. System Architecture

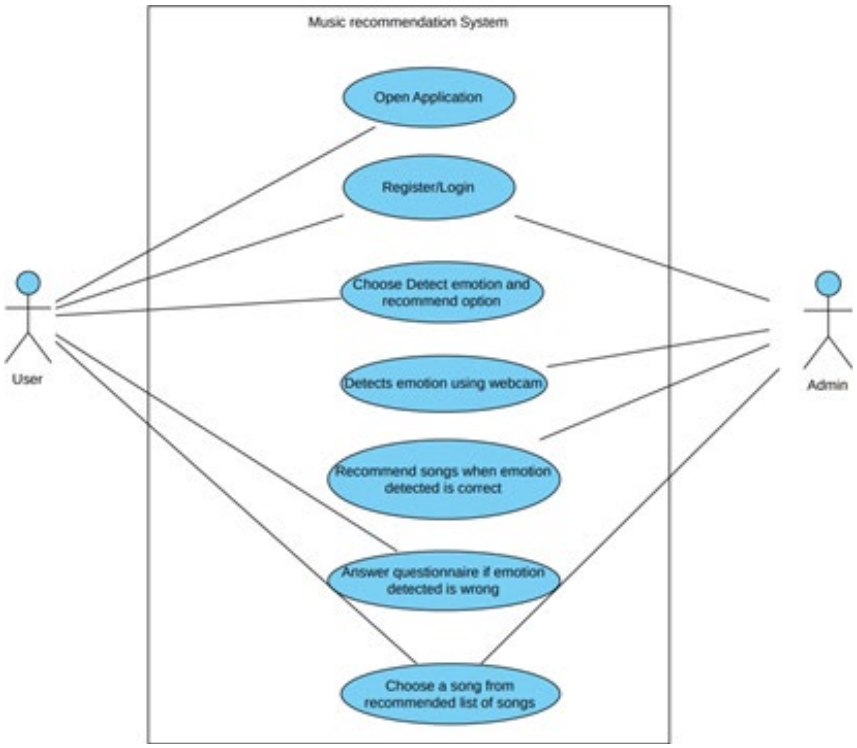

Figure 3: Proposed System Architecture

B. Training and Testing data

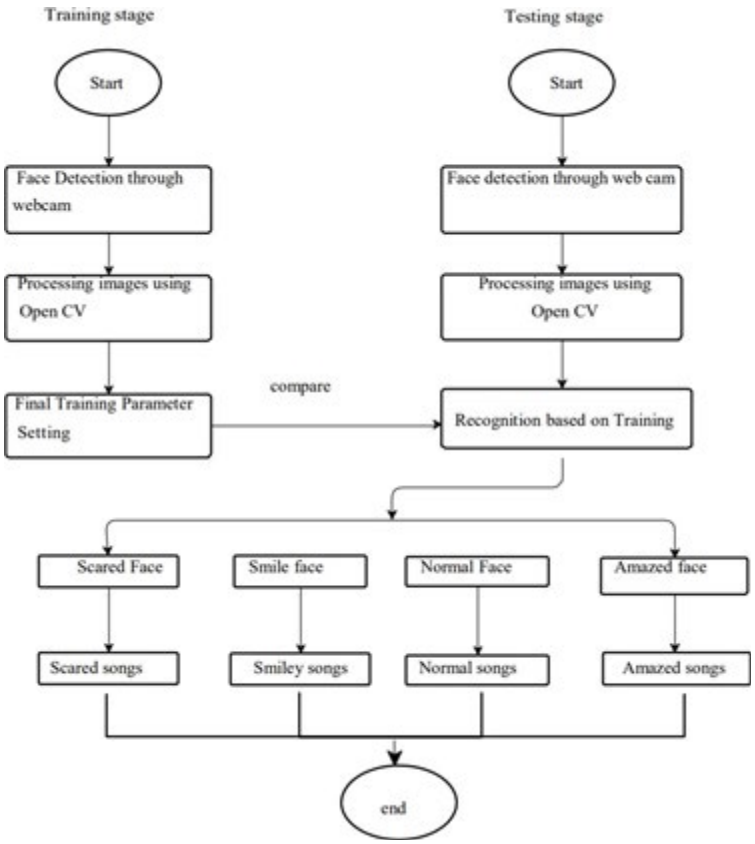

Figure 4: Flow chart

5. Implementation

Implementation includes all conversion activities from the old system to the new system. The old system consists of manual operations that worked very differently from the proposed new system. A good implementation is fundamental to providing a reliable framework to meet the needs of the association. Incorrect settings can affect the results of the electronic framework. Our project is made using the Python programming language. Python provides a compact and easy way to understand the code. Python has a wide range of computational reasoning and AI.

Implementation steps:

1. Open anaconda
2. Select the project environment
3. Start Spyder
4. Run the application app.py
5. Go to Microsoft Edge and enter local host:5000
6. Click register and register
7. Using your registration email, please sign and complete the form existing on the next page.
8. Access the button (Sit up straight and put your face in the middle, otherwise it will appear, please look steadily)
9. After capturing the maximum sentiment, if the error is displayed on the next web page, we also have a questionnaire to know the exact sentiment.

➔ Keras, TensorFlow, and Scikit-learn for AI

➔ NumPy for superior execution of logical registering and information examination

➔ SciPy for cutting-edge figuring

➔ Pandas for broadly useful information examination

➔ Seaborn for information perception

5.1 Convolutional Neural Network: Convolutional neural networks are also called ConvNets / CNNs. The convolutional neural network was developed to overcome the limitation of Artificial neural networks for image-based tasks. In ANN the nodes were arranged column-wise. So, the input to the ANN has to be a 1-dimensional array. A color image has three channels that are red, green, and blue channels. Each channel is a two-dimensional array of pixel values. So, a color image is a three-dimensional array of pixel values. In order to feed to ANN as input we should straighten the three-dimensional array of pixel values into a 1-dimensional array of pixel values. This leads to a loss of spatial information stored in an image. To overcome this, we use CNN.

In CNN rather than arranging the nodes in column-wise, the nodes are arranged in a three-dimensional array. Each layer in CNN is width, height, and depth. Each layer in CNN will be in the form of volume. The output of the previous layers acts as an input for the next layers. In CNN data is stored and processed in the three-dimensional layer. The information through CNN flows in the form of tensors.

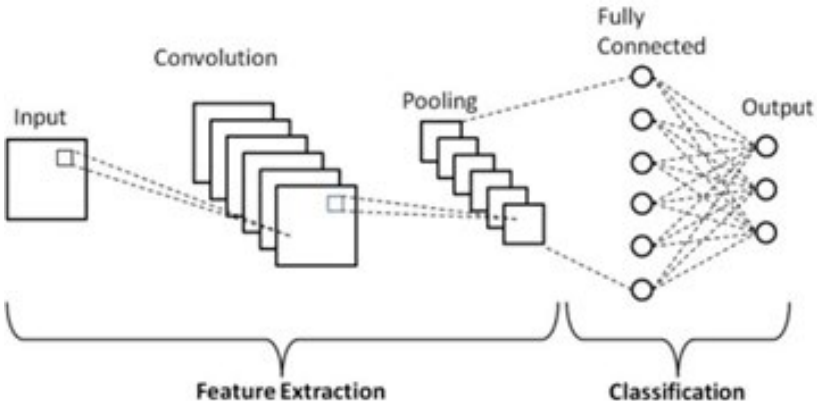

Figure 5: Architecture of CNN

5.2 Data Collection and Data Pre-processing: First we started clicking our own pictures and making a data set, but it was a tedious task for us. So, when we started our research we found that there are many potential data sets that are publicly available. First, we chose a data set that is much smaller and contains images of higher quality, and the pictures in that data set are well-centered which means the position of the eyes, nose, and mouth are similar. On the other hand, we also found the data set which is different in positions. So, for better performance and accuracy, we chose the other dataset. Our data set contains grey-scale images of faces with 48 and the size of the data set is about 57 MB. It contains 7 folders each representing different emotions like surprise, anger, happiness, sadness, neutral, fear, and disgust. There are more than 3000 records.

6. Feature Extraction: Feature extraction is the method for creating a new and smaller set of features that captures most of the useful information of raw data. Feature extraction is the process of copying knowledge from an already trained data set to new data set to solve similar problems. – The main phase of design recognition and information mining is data mining. At this stage, a subset of important elements is extracted from the primary information using explicit guidelines. Large amounts of information can be reduced to smaller sets that are faster to compute. Therefore, qualified item selection is a fundamental step toward effective facial recognition. Features of an image are extracted by its content. Content like color, texture, shape, position, dominant edges of image items, and regions. Image processing is the most important application of feature extraction. Computers see any images in the form of a matrix. The size of the matrix will be depended on the number of pixels of the input image. For every colored image, we will have three matrices also called channels they are Red, Green, and Blue.

7. Emotion Recognition: First the webcam is turned on and using OpenCV face will be detected on capture images and face detection will be pre-processed. This Pre-processed image will be sent to the trained CNN model, the model will classify the emotion into one of the seven categories that are happy, sad, fearful, neutral, surprised, angry, or disgusting. If the user changes his emotion still the emotion will be captured. The face is detected, and the image is pre-processed, the image extracted is suppressed and the features are extracted for further processing then the captured feature is trained classified, and sent to the comparison model, where the emotion is detected from the saved model and compared and finally the emotion is displayed.

8. Spotify Implementation: The request authorization () task handles setting limits for approving claims and getting access tokens. The callAuthorizationApi() job sends a mail request to Spotify account management which returns an access token. This token is used for each request subsequently sent for confirmation. callApi() is used to send a request to Spotify account management to get data such as tracks in a playlist, each currently active device, and accessible audio items for a specific track.

9. Result

S.NO	Input	Output
1	User Registration	Registration successfully
2	User login	Login into the application
3	Click on Capture image or questionnaire	If clicked on the capture option then it captures and detects the emotion. If clicked on the questionnaire it directs you to the set of questions.
4	Check whether the detected emotion is correct or wrong	If emotion is not satisfied, then click on No. Else Yes.
5	If No. Select either capture and recommend or questionnaire.	Same as output 3
6	If clicked on sad emotion	Sad songs will be recommended

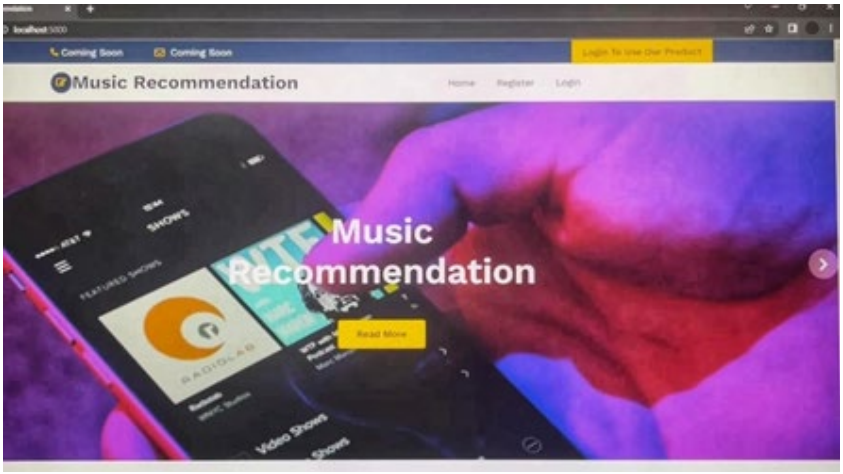

Figure 6: Home page

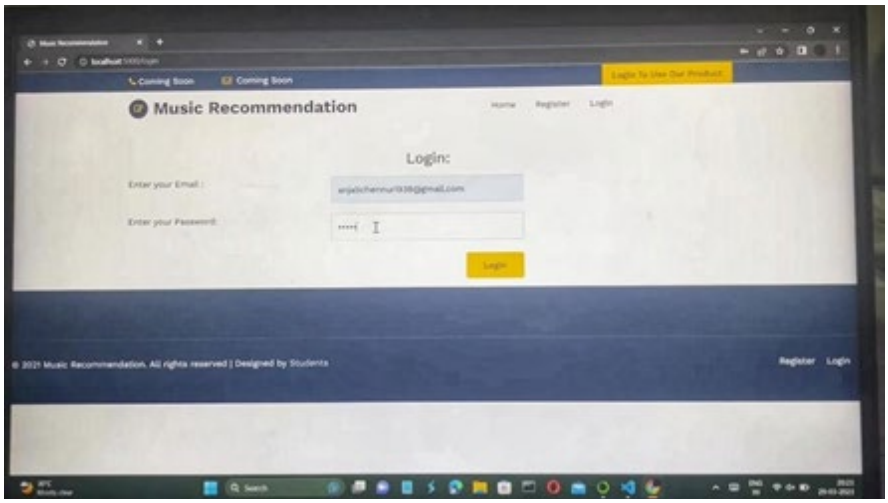

Figure 7: Registration page

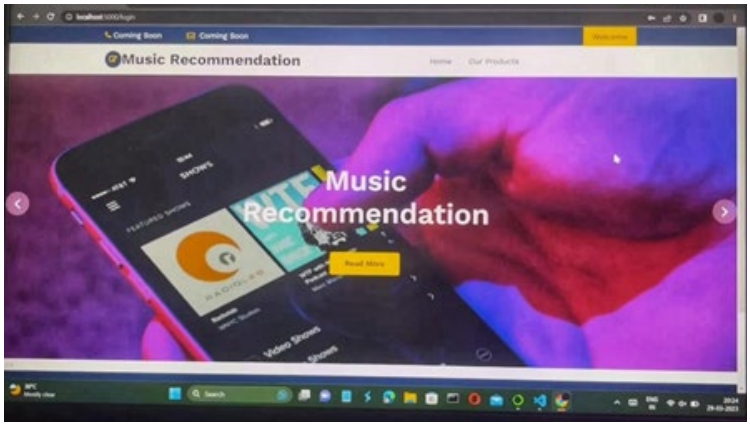

Figure 8: Our Products

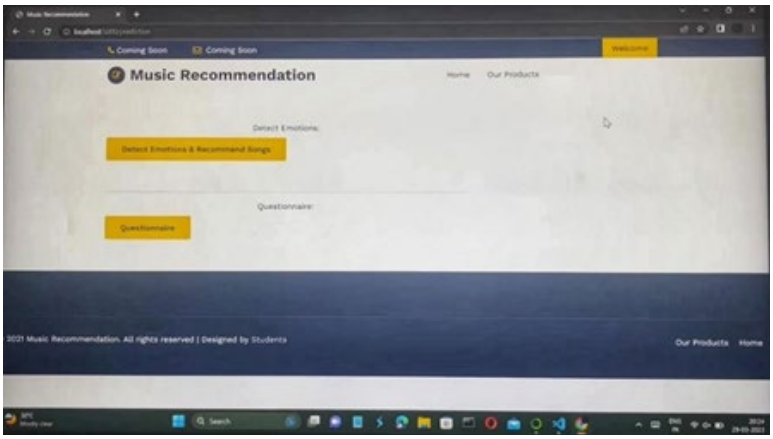

Figure 9: Two models built

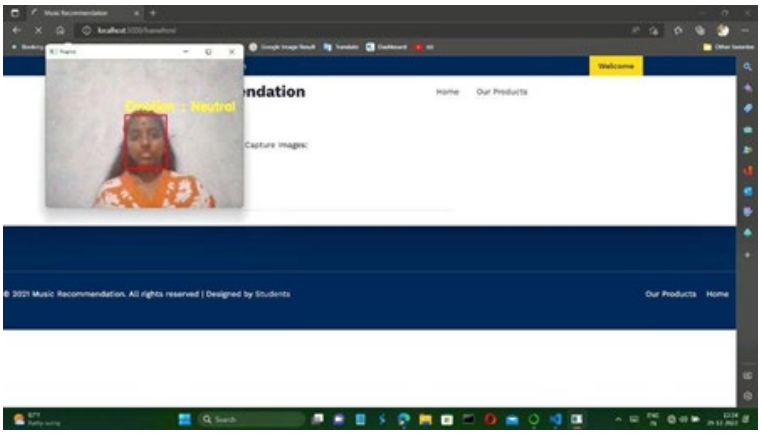

Figure 10: Emotion detection

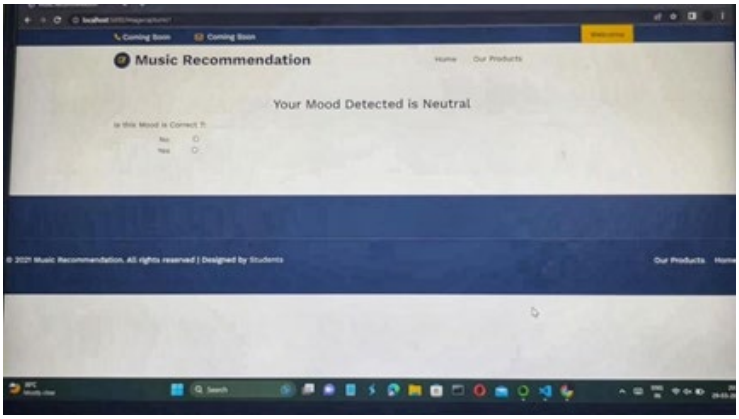

Figure 11: Verifies Whether the emotion is correct or not

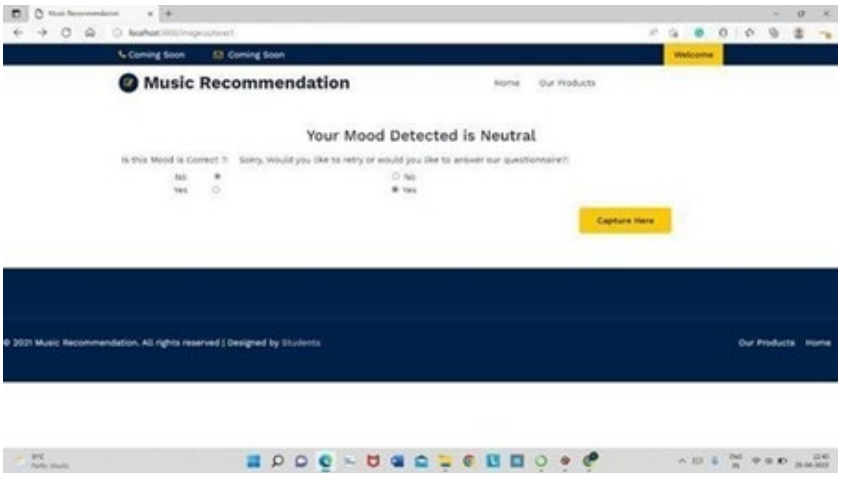

Figure 12: If the user not satisfied with the emotion he can again capture the emotion

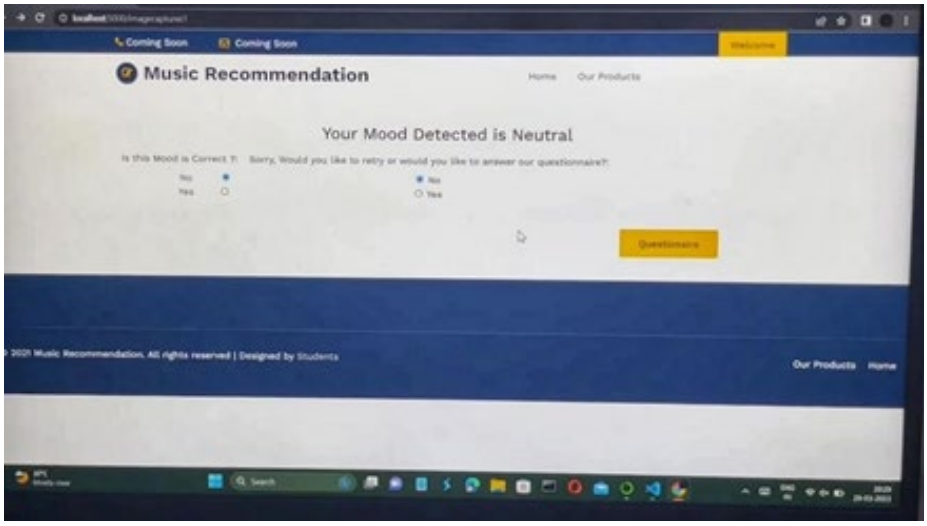

Figure 13: If the user is not satisfied with emotion, he can select any of the models

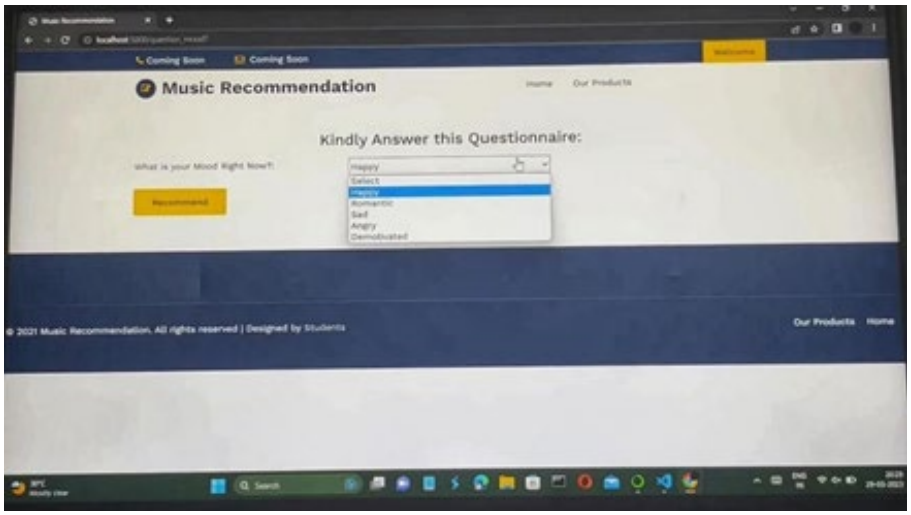

Figure 14: If the questionnaire model is selected user can select any mood manually

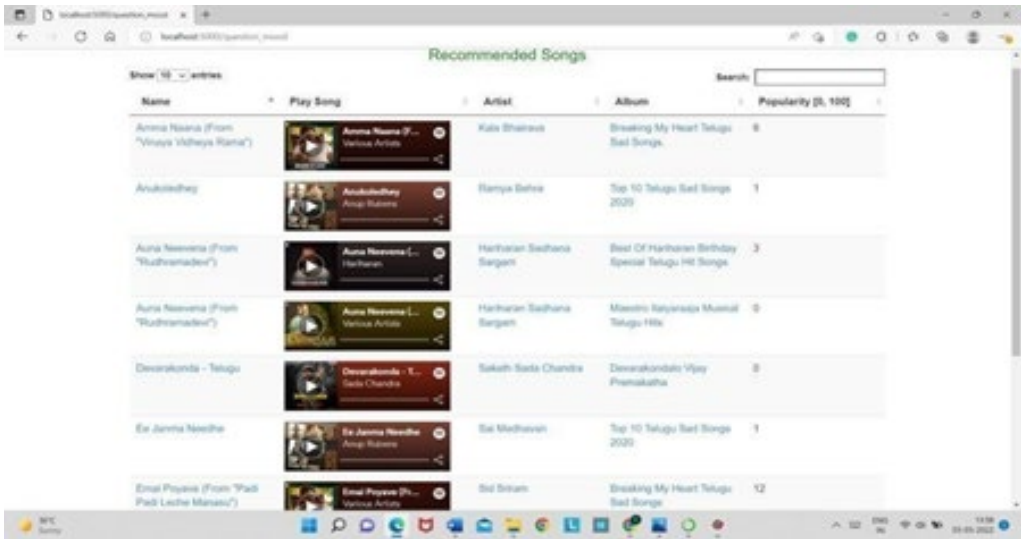

Figure 15: Recommends list of songs

10. Conclusion & Future Scope

The "Music recommendation based on current mood using AI & ML" this system facilitates the use of people by automating and gives the user a better music listening experience. This application resolves the basic needs of music listeners without challenging them as the current existing applications do. The proposed framework basically processes videos of facial behavior, and then plays music based on those emotions. A basic framework for music recommendation using facial emotion recognition is presented here. This recommends music by individual emotions like happiness, anger, sadness, and romance. He can also choose the music according to their state of mind and emotions. The user can access the website at any time. This application increases the user's satisfaction and engages the customers more actively. In the world of fast-paced work culture, this type of system acts as a stress reliever and non-time-consuming application.

The scope of the project can be expanded in the future to enhance its existing functionalities by expanding the model, implementing a more complex model to improve the performance, making the input in speech format, and testing it on a live website.

11. Acknowledgements

We are very grateful to our project guide Mr. Dharmendra K Roy, for his extensive patience and guidance throughout our project work.

We express our intensive gratitude with pleasure to all who have directly or indirectly contributed to our need at the right time for the development and success of our project work.

References

1. Dureha, A. (2014). An accurate algorithm for generating a music playlist based on facial expressions. International Journal of Computer Applications, 100(9), 33-39.
2. Londhe, R. R., & Pawar, V. P. (2012). Analysis of facial expression and recognition based on statistical approach. International Journal of Soft Computing and Engineering, 2(2), 391-394.
3. Gupte, A., Naganarayanan, A., & Krishnan, M. (2016). emotion based music player-XBeats. International Journal of Advanced Engineering Research and Science, 3(9), 236854.
4. Tambe, P., Bagadia, Y., Khalil, T., & Shaikh, N. U. (2015). Advanced Music Player with Integrated Face Recognition Mechanism. undefined.
5. Liu, C., Xie, S., Xie, X., Duan, X., Wang, W., & Obermayer, K. (2018, January). Design of a video feedback SSVEP-BCI system for car control based on improved MUSIC method. In 2018 6th International Conference on Brain-Computer Interface (BCI) (pp. 1-4). IEEE.
6. Cristinacce, D., & Cootes, T. F. (2003, September). Facial feature detection using AdaBoost with shape constraints. In BMVC (pp. 1-10).
7. Lienhart, R., & Maydt, J. (2002, September). An extended set of haar-like features for rapid object detection. In Proceedings. international conference on image processing (Vol. 1, pp. I-I). IEEE.
8. Menezes, P., Barreto, J. C., & Dias, J. (2004). Face tracking based on haar-like features and eigenfaces. IFAC Proceedings Volumes, 37(8), 304-309.
9. Muller, N., Magaia, L., & Herbst, B. M. (2004). Singular value decomposition, eigenfaces, and 3D reconstructions. SIAM review, 46(3), 518-545.
10. Wilson, P. I., & Fernandez, J. D. (2022, May). Establishing a face recognition research environment using open source software. In 2005 GSW.
11. Hemanth P,Adarsh ,Aswani C.B, Ajith P, Veena A Kumar, EMO PLAYER: "Emotion Based Music Player," International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 4, April 2018, pp. 4822-87.
12. Happy, S. L., & Routray, A. (2014). Automatic facial expression recognition using features of salient facial patches. IEEE transactions on Affective Computing, 6(1), 1-12.
13. Su, J. H., Liao, Y. W., Wu, H. Y., & Zhao, Y. W. (2020, October). Ubiquitous music retrieval by context-brain awareness techniques. In 2020 ieee international conference on systems, man, and cybernetics (smc) (pp. 4140-4145). IEEE.
14. Vivek, J. D. R., Gokilavani, A., Kavitha, S., Lakshmanan, S., & Karthik, S. (2017, March). A novel emotion recognition

based mind and soul-relaxing system. In 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS) (pp. 1-5). IEEE.
15. Liu, C., Xie, S., Xie, X., Duan, X., Wang, W., & Obermayer, K. (2018, January). Design of a video feedback SSVEP-BCI system for car control based on improved MUSIC method. In 2018 6th International Conference on Brain-Computer Interface (BCI) (pp. 1-4). IEEE.
16. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 ieee computer society conference on computer vision and pattern recognition-workshops (pp. 94-101). IEEE.
17. Hadid, A., Pietikäinen, M., & Li, S. Z. (2007). Learning personal specific facial dynamics for face recognition from videos. In Analysis and Modeling of Faces and Gestures: Third International Workshop, AMFG 2007 Rio de Janeiro, Brazil, October 20, 2007 Proceedings 3 (pp. 1-15). Springer Berlin Heidelberg.
18. Dhavalikar, A. S., & Kulkarni, R. K. (2014, February). Face detection and facial expression recognition system. In 2014 International Conference on Electronics and Communication Systems (ICECS) (pp. 1-7). IEEE.
19. Lee, Y. H., Han, W., & Kim, Y. (2013, September). Emotional recognition from facial expression analysis using bezier curve fitting. In 2013 16th International Conference on Network-Based Information Systems (pp. 250-254). IEEE.
20. Londhe, R. R., & Pawar, V. P. (2012). Analysis of facial expression and recognition based on statistical approach. International Journal of Soft Computing and Engineering, 2(2), 391-394.
21. Rahman, J. S., Gedeon, T., Caldwell, S., Jones, R., & Jin, Z. (2021). Towards effective music therapy for mental health care using machine learning tools: human affective reasoning and music genres. Journal of Artificial Intelligence and Soft Computing Research, 11(1), 5-20.