International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Music Mood Based Recognition System Based on Machine Learning and Deep Learning

Dr. R. Priscilla Joy[1], Dr. M. Roshni Thanka[2], Dr. Sangeetha[3], Dr. Julia Punitha Malar Dhas[*] [4], Dr. E. Bijolin Edwin[5,] Dr. Ebenezer[6]

Submitted: 15/11/2022 **Revised:** 11/01/2023 **Accepted:** 10/02/2023

Abstract*:* There are extensive studies about music's impact on human's emotional state. Humans detect a wide range of emotions from various genres of music, and music plays an integral role in personality development and the treatment of ailments. Music has tremendous effects on human moods and thoughts. Consequently, it impacts cognitive and biological health, and the concept of well-being through music is acquiring traction. In the treatment of depression, music therapy gets witnessed as an addendum to psychoanalysis. Music can enhance intellectual and physical work, study, sports, relaxation, relieve fatigue, and music therapy, among other things. People often get confused while searching for music according to their interests and mood. Individuals usually listen to a particular genre or performer when they are in a certain mood. Music has the ability to control mood, specifically to boost energy, and reduce anxiety. Listening to the correct song at the opportune timing, may help with mental health. As a result, human mood changes and music have an interdependent affinity. In this paper, we aim to develop an application that can understand facial features (Mood and Emotions) and recommend music accordingly using Machine Learning and Deep Learning as tools and algorithms.

Keywords: *Machine Learning, Deep Learning, OpenCV, Music, Facial Recognition, Mood.*

1. Introduction

In this system, we propose a design of a personalized emotion-driven music recommendation system. The main goal is to solve the choice dilemma, discover new music, support emotional and physical well-being, and aid in the improvement of the song selection process. Using a combination of artificial intelligence algorithms and generic music recommendations using facial feature recognition methodologies are to be used in the design. Our personalized song recommendation system is a prototype that focuses on mood recognition in an instant and suggests songs accordingly. The prototype of our product/API consists of two main modules:

- Music recommendation based on artists
- Facial expression recognition/mood detection

Using "Valence-Arousal Plane" alongside vector-distance measurements using K-Means clustering algorithm to match tracks that convey a similar type of emotion/mood.

1.1 Music Recommendation Model

James Russell [1] proposed the circumplex concept of emotion in 1980. According to this concept, emotions get dispersed in a circular 2D space with arousal and valence dimensions. The horizontal x-axis indicates valence, and the vertical y-axis indicates arousal, with a balanced value of valence or a moderate level of arousal represented by the center of the circle. In this approach, emotional variables are expressed at various valence and arousal levels, as well as at the neutral/non-emotional level of one or both of these components. Most typically, circumplex models are used to assess emotion words, emotional facial expressions, and effective states. Our model is based upon the concept of the Valence-Arousal Plane. The following diagram shows the circumplex model of the plane. This theory is used as a benchmark to classify the moods in any type of music. The figure also contains the numerical values needed to classify the emotion taken from Patrick Helmholz's paper. [2]

1 Karunya Institute of Technology and Sciences, Coimbatore
ORCID ID : 0000-0001-5778-8415
2 Karunya Institute of Technology and Sciences, Coimbatore
ORCID ID : 0000-0002-8129-3663
3Karunya Institute of Technology and Sciences, Coimbatore
ORCID ID : 0000-0003-2708-231X
4 Karunya Institute of Technology and Sciences, Coimbatore
ORCID ID: 0000-0001-5563-8077
5 Karunya Institute of Technology and Sciences, Coimbatore
ORCID ID : 0000-0003-3904-1902
**Corresponding Author Email: juliapunitha@karunya.edu*

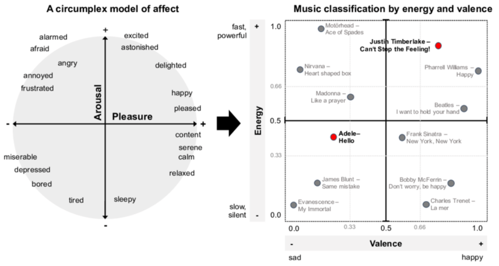


Fig 1. Classification of songs through energy and valence graph

From the figure 1 it's clear that valence and energy values fall in the range between 0.0 to 1.0 with four base classes of emotion: *Happy, Calm, Sad, Angry, and Neutral as the origin.* Here Classifying songs through energy and valence graph based on the human emotions-Model by Russell. [2]

The quadrants of the model that result are as follows in table 1:

Table 1: Classifying correlating valence and energy values to form our base quadrants.

Quadrant	**1**	**2**	**3**	**4**
Emotion	Happy	Calm	Sad	Anger
Valence	0.5 to 1.0	0.5 to 1.0	0 to 0.5	0 to 0.5
Energy	0.5 to 1.0	0 to 0.5	0 to 0.5	0.5 to 1.0

1.2 Emotion Recognition Theory

The technology-based methodology of identifying a human face is facial recognition. Biometrics is used in a facial recognition system to map facial digital images or videos. To get a match, it will compare the data to a dataset of fed faces. Facial recognition can aid in the verification of a person's identity, but it also raises privacy concerns. One of the face recognition technologies that has evolved and increased over time is emotion recognition. Currently, facial emotion recognition software is utilized to allow a program to inspect and process a human's facial expressions. This program, which uses complex image dispensation to act like a human brain, is also capable of identifying emotions. It is AI, or "Artificial Intelligence," that recognizes and analyses various facial expressions to combine them with other data. This can be used for a multitude of purposes, such as investigations and interviews, and allows authorities to identify a person's emotions/moods using just technology.

1.3 Song Database

Spotipy is a light Python library for Spotify Web API. With Spotipy, one gets a complete pass to the entire music data supplied by the Spotify platform. Spotipy backs up all of the components of the Spotify Web API, including access to endpoints plus aid for user's authorization. Authorization is via the Spotify Accounts service, and each user who signs up on the Developer platform gets one Client User ID and one Client Secret ID for license purposes which will link our code to the Spotify server platform directly to extract the data accordingly, to the code we've written. We wrote a solution to web-scrape the required data needed directly from Spotify using Python and Jupyter/Spyder/VSCode as the IDE. After registering for Spotify Developers, and getting approved, we generated the Spotify Client User ID and Client Secret ID. Using these particular credentials. we can generate a customized artist-based playlist. We web scrape the data generated into a CSV file with the following columns: song id, genre, artist name, track name, valence values, and energy values. We now manipulate the dataset according to our needs. We performed exploratory data analysis on our dataset. Moreover, we added more attributes which help us analyse and gain more insights about the music tracks such as danceability, loudness, tempo, speechiness, instrumentalness and popularity.

1.4 Using Valence-Arousal Theory for K-Means Clustering

This dataset is stored as a CSV file and data preparation begins with us using the Valence-Arousal theory to classify the song database into the respective mood quadrants.

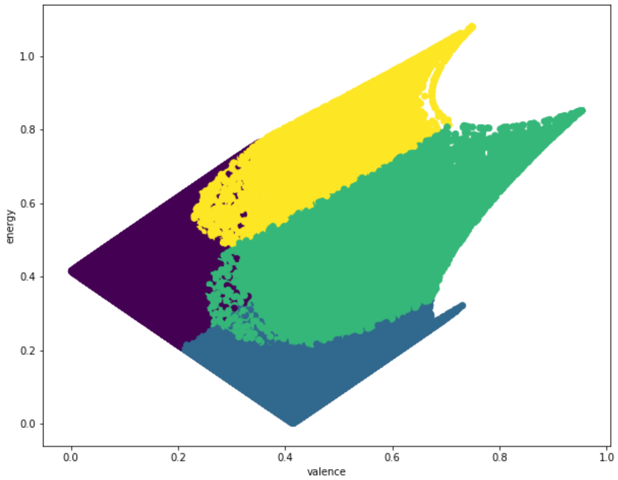


Fig 2. Valence-Energy graph based on K-Means clusters on the song database.

As we've seen in the James Russell [1] and Thayer's model [3], there are a total of twenty moods, i.e., four base emotions with four more subclasses in each quadrant. Resulting in five emotions per quadrant as in fig. 2. In correspondence to this, we use our four base clusters and corresponding values to create a mood vector, a 2D array that will contain the values for each song/track the valence and energy values.

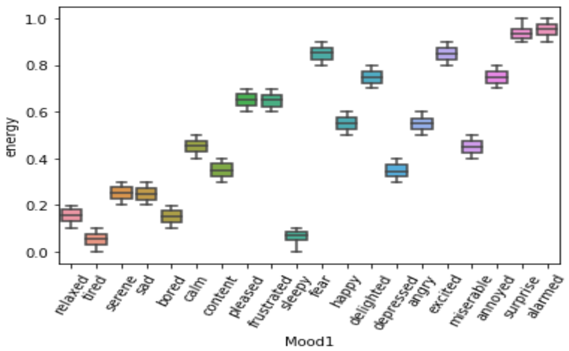


Fig. 3. Energy vs. Subclassified Mood Boxplot

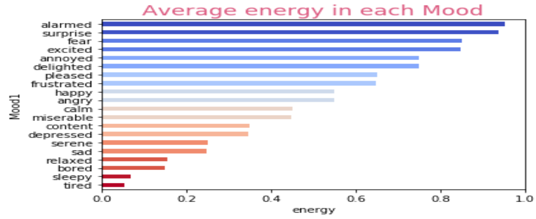


Fig. 4. Bar chart for average energy in each Mood

As we analyse our data through boxplots as in fig.3 and bar charts in Fig. 4, we realize that each sub-classed emotion can be represented with a 0.1 difference in terms of the hierarchy of emotions using the energy values of each particular song in the database.

Emotion Recognition Module- Emotion Classifier Selection

Now we need to incorporate the facial recognition of the user using OpenCV and Keras. To select our algorithm, we analysed the papers [4]-[6] as tabulated in table 2 below

Table 2: Observations on Algorithms

Algorithm	***Dataset Used***	***Description***	***Observation***
Fisherface algorithm	CK extended + HELEN data set	Uses a linear combination of features that maximize the total variance in data.	It doesn't consider any classes and so a lot of discriminative info or features may be lost. This model only detected 2 emotions with 80% accuracy.
Eigenfaces + SVM [5]	Million Song Dataset by Kaggle	Uses static imaging and statistical approach to defect large pose variations with particular feature selection.	Works well only when applied with machine learning algorithms and requires particular feature selection with good camera resolution.
Haar cascade/ Viola-Jones algorithm [6]	Various data sets	Requires grayscale pictures. The intensity of gray will be used to detect features. Can be used for both static and real-time imaging.	Works well with all algorithms and is largely used. Is readily available for low-res. cameras too.

Based on the above tables arrived with the following requirements:

- Static and real-time imaging
- Works well with deep learning and machine learning models
- Is readily available
- Works for low and high-resolution cameras

Out of all the three classifiers, we found that Haar Cascade/Viola-Jones Framework to be the most suitable for our needs as it matches all our requirements.

2. Related Works

The mood-based music system is a piece of computer software that focuses on mood recognition. It's a concept for a new product that combines several different interfaces:

face detection, facial expression identification, song mood classification, appropriate recommended music, and playlist building. Our model’s detailed architecture as per progress and projections are depicted in Fig. 5.

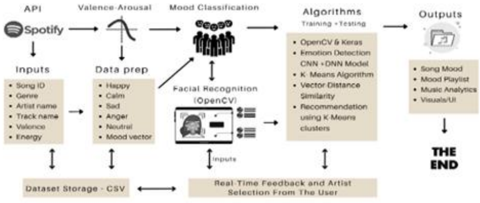


Fig. 5. Architecture

2.1 Deep Convolutional Neural Network

Machine Learning is a critical piece of Artificial Intelligence where one can feed data to the system or the machine to make the machine understand and learn processes and functions through patterns. The system can now predict and provide solutions for like futuristic problems. Neural Network, in general terms, is based on the working of the human brain and works like the same. Meanwhile, if our problems are associated with images, it comes under a field of computing called Computer Vision. During this paper, we've got to leverage some concepts of Machine Learning and Deep Learning. Beginning from reading and manipulating the datasets to developing models to acknowledge facial expression. Deep Convolutional Neural Network is one such concept we've employed in our paper. Convolution is merely a mathematical mix of functions supplying another different function. Also, uniting two clusters of knowledge. Concerning the subject of Convolutional Neural Networks or CNN, convolution gets performed on an input file incorporating a kernel or maybe a filter to reduce some feature map. Convolutional neural networks are usually used for visual imagery, helping the personal computer identify and learn from images. However, when a neural network, CNN per se, usually of two or more layers and a suitable level of complexity entitles to a Deep Neural Network or DNN. Our DCNN model has six convolution layers, seven batch normalization layers, three max-pooling layers, four dropout layers, one flattens and dense layer with an input and output layer. Other than that, to boost our model's optimization, we've used a Nadam optimizer. Our DCNN model has been trained for approximately 60 epochs with an accuracy of 76% to 78% classified with four major emotions using the FERPlus dataset. While developing the model of the Facial Recognition system, we have deployed the below listed layers of the Deep Convolutional Neural Network.

1. **Input Layer:** The image data should be stored in the input layer of a CNN. As we saw before, image data can be represented by a three-dimensional matrix. You must resize it to fit into a single column. If you have a 28 × 28 = 784-pixel image, you'll need to transform it to 784 x 1 prior to actually feeding it into the input. The dimensionality of input would be "m" if you have "m" training samples (784, m).

2. **Convolutional Layer/Conv2D:** The layer yields a tensor of outcomes by convolving a kernel/filter with the layer input. The bias vectors are constructed and appended to the outputs if use bias is True. Lastly, when activation is not None, the outputs are activated as well. Also called the COV2D Layer.

3. **Pooling Layer:** Pooling is now a technique for reducing image size. We have used Max Pooling. It's nothing more than determining the most heightened value from a matrix of a certain dimension (the default size for it is 2 X 2). This method is useful for extracting important characteristics or features that are highlighted in the image.

4. **Batch Normalization:** A layer that allows each layer of the network to learn more independently is batch normalization. It's used to make the output of the previous layers more normal. The activations, in this normalization, adjust the input layers. Training the data through this batch normalization is way easy and efficient. It can also be used to avoid model overfitting by regularization. In order to standardize the intakes and results, this gets applied to a sequential model.

5. **Dropout Layer:** A Dropout layer is another common characteristic of CNNs. The Dropout layer is like a masking that cancels out some neurons' contributions to the following layer while leaving all others unchanged. We can use a Dropout layer to nullify some of the features of an input vector, but we can also use it to nullify some hidden neurons in a hidden layer. Dropout layers are crucial in CNN training because they prevent the training data from being overfit. If they're missing, the very first batch of training sets has a disproportionately large impact on learning.

6. **Flattening Layer:** The process of converting data into a two-dimensional array for subsequent processing is known as flattening. A single long feature vector is created from the outputs of the flattened convolutional layers. Then it's further connected to the final classifier, resulting in a fully-connected layer. In other words, the combination of all the pixel data compiled into a single line, connected to the final layer.

7. **Dense Layer:** The dense layer is a basic layer containing neurons, with each neuron receiving input from all the NEURONS in the layer before it. The Dense Layer takes the results of convolutional layers to classify images.

3. System Analysis

3.1 Dataset Selection for Facial Recognition

We now need to create our facial recognition model that we can now use on top of our classifier that is the Haar Cascade classifier. As we surveyed various datasets for this research, we came across a few datasets namely: FER 2013[7], CK (Cohn- Kanade) & CK+(Extended) [8], KDEF (Karolinska Directed Emotional Faces)[9], and FERPlus[10]. Although every dataset has its advantages, we moved ahead with FERPlus. This dataset successfully satisfies our requirements involved in this paper, which are:

- Diverse emotions
- Higher accuracy for the facial detector model
- Near about four emotions classes
- Large training and testing set

The FER Plus dataset comprises of:

- Monochromatic images of faces at a resolution of 48x48 pixels.
- The dataset includes eight categories of emotions.

We'll be using four classes for this paper, namely: Angry, Happy, Sad, and Neutral/Calm.

So, the dataset has a total of 24256 photos for training and 3006 images for validation. In [11]-[17] it is explained and analyzed how the music and mood interrelated with physcology. [18]-[23] it is explaining about the music mood and the implementation using machine learning.

3.2 Trials and Testing

We trained our model through a series of trials and tests. The results of the same are specified in the table 3 below.

Table 3. Trials

TRIALS	EMOTIONS	ACCURACY	EPOCHS
Simple CNN	6 Emotions	Below 65	40
	4 Emotions	Above 70	55
Deep CNN	6 Emotions	Below 72	45
	4 Emotions	About 80	60

We used Deep Convolutional Neural Network while testing our model with different emotions and number of epochs. We learnt that fewer emotion classes, with greater number of epochs and more complex Convolutional Neural Network can provide better accuracy as compared to the rest. As we know that a model is considered good only when the data lost while running and accuracy of the model is optimized. We have generated two graphs that can visually demonstrate how good this model is. Let us discuss these graphs below.

Graph 1: Loss for Emotion Model: We have generated an epoch against loss graph for both training and validation data. We can clearly see that as the number of epochs increases the loss in data for the model decreases.

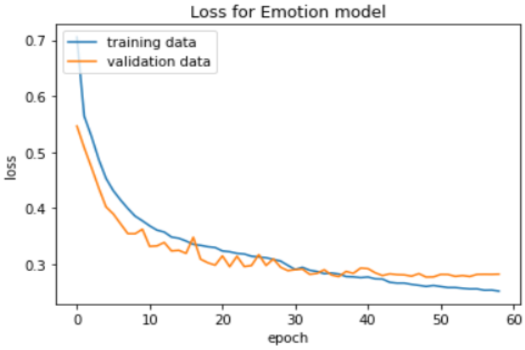


Fig. 6. Loss Graph for our Emotion model

The loss percentage initially is over 60% and 50% for training and validation data respectively. But as we go further in the graph, the value for loss decreases to 30% or below for both the data, provided the number of epochs is greater. The Loss graph and accuracy for emotion model is depicted in fig. 6 and Fig. 7 respectively.

Graph 2: Accuracy for Emotion model: We generated another graph for epoch against accuracy for our model for our model for both training and test data. It is evident that as the number of epochs increases, the graph has an exponential increase in the accuracy. The accuracy is 30% and 40% at zeroth epoch for train and test data respectively. But it rises to about 65% in the first 10 epochs and gradually increases with the number of epochs.

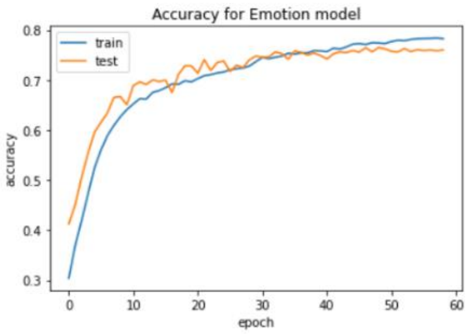
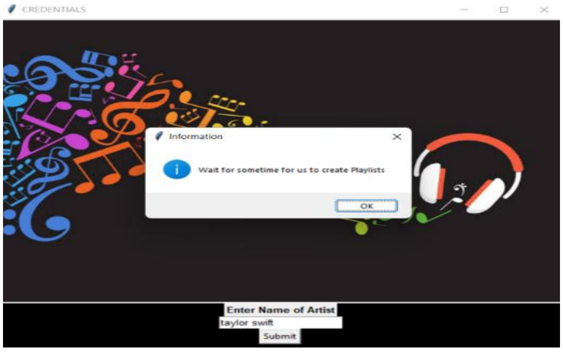


Fig. 7. Accuracy Graph for our Emotion model

4. Results and Discussion

As we’ve created our Deep Convolution Neural Network the next step for us is to create our recommendation engine using the concepts of Vector-Distance similarity and K-Means Clustering Algorithm for the songs in our final mood dataset. Our aim is to create a playlist that after getting the valid output from the facial recognition side, gives the user an option to delve deeper into that emotion or try to move towards a positive side. For example: If a user is sad/depressed, the user will be given a playlist option that has valence-energy values correlating to the upper side of the quadrant/graph so as to cheer up the user into a better emotion like happiness or calm/serene.



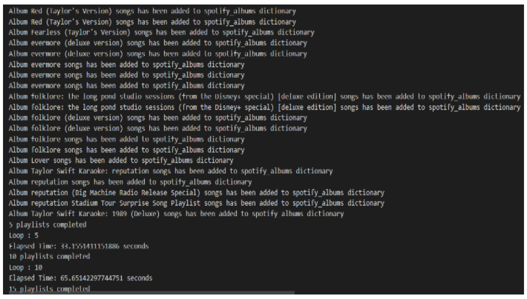
Wait for sometime for us to create Playlists

Enter Name of Artist: taylor swift

Fig. 8. First Window: Enter Artist Name

We’ve used Tkinter to render our GUI, our first window asks the user to type an artist name according to their choice and taste, followed by a prompt to the user to wait a while. Figure 8 represents the same.

The system then moves on to extract live data from Spotify using the API, Spotipy and generates a live database of songs and then proceeds to make playlists based on the user’s choice. Figure 9 represents the same.



```
Album Red (Taylor's Version) songs has been added to spotify_albums dictionary
Album Red (Taylor's Version) songs has been added to spotify_albums dictionary
Album Fearless (Taylor's Version) songs has been added to spotify_albums dictionary
Album evermore (deluxe version) songs has been added to spotify_albums dictionary
Album evermore (deluxe version) songs has been added to spotify_albums dictionary
Album evermore songs has been added to spotify_albums dictionary
Album evermore songs has been added to spotify_albums dictionary
Album evermore songs has been added to spotify_albums dictionary
Album folklore: the long pond studio sessions (from the Disney+ special) [deluxe edition] songs has been added to spotify_albums dictionary
Album folklore: the long pond studio sessions (from the Disney+ special) [deluxe edition] songs has been added to spotify_albums dictionary
Album folklore (deluxe version) songs has been added to spotify_albums dictionary
Album folklore (deluxe version) songs has been added to spotify_albums dictionary
Album folklore songs has been added to spotify_albums dictionary
Album folklore songs has been added to spotify_albums dictionary
Album Lover songs has been added to spotify_albums dictionary
Album Taylor Swift Karaoke: reputation songs has been added to spotify_albums dictionary
Album reputation songs has been added to spotify_albums dictionary
Album reputation (Big Machine Radio Release Special) songs has been added to spotify_albums dictionary
Album reputation Stadium Tour Surprise Song Playlist songs has been added to spotify_albums dictionary
Album Taylor Swift Karaoke: 1989 (Deluxe) songs has been added to spotify_albums dictionary
5 playlists completed
Loop : 5
Elapsed Time: 33.1551411151886 seconds
10 playlists completed
Loop : 10
Elapsed Time: 65.65142297744751 seconds
15 playlists completed
```

Fig. 9. Extracting live data from Spotify and generating a playlist

After the generation of full artist playlists, our OpenCV model comes into play and the user’s live emotion is captured. To capture the correct emotion the user can press ENTER key repeatedly until the desired emotion is on the screen and press Q to quit the video feed. Our Emotion Recognizer model takes the live video feed and derives emotion from an accurate snapshot while cropping the image's region of interest, i.e. the facial area from the image which is then a grayscale image, moved to processing resulting in the emotion required. The live emotion recognition is given in fig. 10.

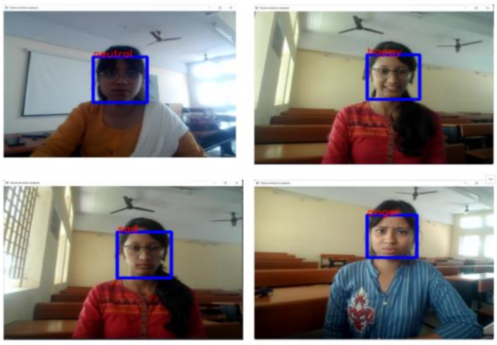
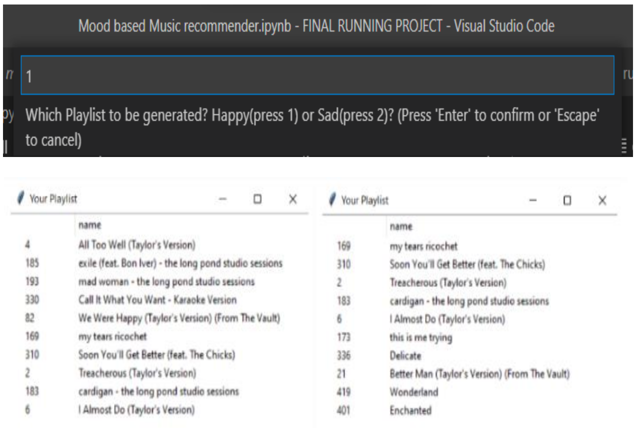


Fig. 10. Live Emotion Recognition

As the mood is detected we press Q, the program moves to the terminal. In the terminal the algorithm prints the best few songs’ playlist that is most likely to be heard by the user.



Mood based Music recommender.ipynb - FINAL RUNNING PROJECT - Visual Studio Code

1

Which Playlist to be generated? Happy(press 1) or Sad(press 2)? (Press 'Enter' to confirm or 'Escape' to cancel)

Your Playlist

	name
4	All Too Well (Taylor's Version)
185	exile (feat. Bon Iver) - the long pond studio sessions
193	mad woman - the long pond studio sessions
330	Call It What You Want - Karaoke Version
82	We Were Happy (Taylor's Version) (From The Vault)
169	my tears ricochet
310	Soon You'll Get Better (feat. The Chicks)
2	Treacherous (Taylor's Version)
183	cardigan - the long pond studio sessions
6	I Almost Do (Taylor's Version)

Your Playlist

	name
169	my tears ricochet
310	Soon You'll Get Better (feat. The Chicks)
2	Treacherous (Taylor's Version)
183	cardigan - the long pond studio sessions
6	I Almost Do (Taylor's Version)
173	this is me trying
336	Delicate
21	Better Man (Taylor's Version) (From The Vault)
419	Wonderland
401	Enchanted

Fig. 11. User’s Choice for Mood Playlist, User’s Mood Playlist (Happy Playlist for Sad Emotion)

For emotions like Anger and Sadness people tend to listen to a playlist that can uplift their mood or sometimes stay in the same mood. For such use cases, we have added an option to ask the user if they want a mood uplift or not. For Anger emotion, options are given as Calm or Anger and for Sad emotion, options for Happy or Sad. As per the user’s choice the playlist gets generated in figure 11.

Our final aim is to accurately describe the mood of the user, songs in our database as well as the live data that is extracted. Then, create a mood playlist and update it in real-time and provide options to the user as per their mood to uplift or stay in the same emotion. Hence to, analyse the music taste and render the application using appropriate visuals and UI.

5. Conclusion

Music can be used for a variety of purposes, including supporting behavioral and cognitive work, education, athletics, relaxation, anxiety and fatigue reduction, music therapy, and so on. This proposed model aims to give us a higher accuracy with recommendations for facial feature detection while generating a real-time playlist according to mood and suggesting mood improvement through AI song selection. Hence, we conclude this method can be said to have a better accuracy than genre classification and can be used to improve social networking & customer satisfaction.

References

[1] Russell, James. (1980) “A Circumplex Model of Affect,” Journal of Personality and Social Psychology. 39. 1161-1178. 10.1037/h0077714.

[2] Helmholz, Patrick & Meyer, Michael & Robra-Bissantz, Susanne. (2019)” Feel the Moosic: Emotion-based Music Selection and Recommendation” 10.18690/978-961-286-280-0.11.

[3] Nguyen, Van & Kim, Donglim & Ho, V.P. & Lim, Younghwan. (2017)”A New Recognition Method for Visualizing Music Emotion,” International Journal of Electrical and Computer Engineering. 7. 1246-1254. 10.11591/ijece.v7i3.pp1246-1254.K.

[4] Mustamin Anggo and La Arapu 2018 J. Phys.: Conf. Ser. 1028 012119

[5] Deny John Samuvel, B. Perumal and Muthukumaran Elangovan, "Music rec-ommendation system based on facial emotion recognition", 2020.

[6] K. Vikram and S. Padmavathi, "Facial parts detection using Viola Jones algorithm," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017, pp. 1-4, doi: 10.1109/ICACCS.2017.8014636.

[7] Giannopoulos, P., Perikos, I., Hatzilygeroudis, I. (2018). Deep Learning Approaches for Facial Emotion Recognition: A Case Study on FER-2013. In: Hatzilygeroudis, I., Palade, V. (eds) Advances in Hybridization of Intelligent Methods. Smart Innovation, Systems and Technologies, vol 85. Springer, Cham. https://doi.org/10.1007/978-3-319-66790-4_1

[8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

[9] Ellen Goeleven, Rudi De Raedt, Lemke Leyman & Bruno Verschuere (2008) “The Karolinska Directed Emotional Faces: A validation study, Cognition and Emotion,” 22:6, 1094-1118, DOI: 10.1080/02699930701626582

[10] Stewart Joanna, Garrido Sandra, Hense Cherry, McFerran Katrina, “Music Use for Mood Regulation: Self-Awareness and Conscious Listening Choices in Young People with Tendencies to Depression” JOURNAL=Frontiers in Psychology, VOLUME=10, YEAR=2019, DOI=10.3389/fpsyg.2019.01199, ISSN=1664-1078

[11] Ahmad, Nawaz & Rana, Afsheen. (2015). Impact of Music on Mood: Empirical Investigation. Research on Humanities and Social Sciences. 5. 98-101.

[12] McCraty, Rollin & Barrios-Choplin, B & Atkinson, M & Tomasino, Dana. (1998). The effects of different types of music on mood, tension, and mental clarity. Alternative therapies in health and medicine. 4. 75-84.

[13] Ahmad, Nawaz and Rana, Afsheen, Impact of Music on Mood: Empirical Investigation (November 29, 2015). Research on Humanities and Social Sciences. ISSN (Paper) 2224-5766 ISSN (Online) 2225-0484 (Online), Available at SSRN: https://ssrn.com/abstract=2696883

[14] Stewart J, Garrido S, Hense C, McFerran K. Music Use for Mood Regulation: Self-Awareness and Conscious Listening Choices in Young People With Tendencies to Depression. Front Psychol. 2019 May 24;10:1199. doi: 10.3389/fpsyg.2019.01199. PMID: 31178806; PMCID: PMC6542982.

[15] A. Baharum, T. W. Seong, N. H. M. Zain, N. M. M. Yusop, M. Omar and N. M. Rusli, "Releasing stress using music mood application: DeMuse," 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2017, pp. 351-355, doi: 10.1109/ICTC.2017.8191001.

[16] Xue, H., Xue, L., Su, F. (2015). Multimodal Music Mood Classification by Fusion of Audio and Lyrics. In: He, X., Luo, S., Tao, D., Xu, C., Yang, J., Hasan, M.A. (eds) MultiMedia Modeling. MMM 2015. Lecture Notes in Computer Science, vol 8936. Springer, Cham. https://doi.org/10.1007/978-3-319-14442-9_3

[17] Campbell, E. A., Berezina, E., & Gill, C. M. H. D. (2021). The effects of music induction on mood and affect in an Asian context. Psychology of Music, 49(5), 1132–1144. https://doi.org/10.1177/0305735620928578

[18] Miguel Civit, Javier Civit-Masot, Francisco Cuadrado, Maria J. Escalona,A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends,Expert Systems with Applications,Volume 209,2022,118190,ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.118190.

[19] Garg, Anupam & Chaturvedi, Vybhav & Dhindsa, Arman Beer & Varshney, Vedansh & Parashar, Anshu. (2022). Machine learning model for mapping of music mood and human emotion based on physiological signals. Multimedia Tools and Applications. 81. 10.1007/s11042-021-11650-0.

[20] S. Deebika, K. A. Indira and Jesline, "A Machine Learning Based Music Player by Detecting

Emotions," 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2019, pp. 196-200, doi: 10.1109/ICONSTEM.2019.8918890.

[21] Han, D., Kong, Y., Han, J. et al. A survey of music emotion recognition. Front. Comput. Sci. 16, 166335 (2022). https://doi.org/10.1007/s11704-021-0569-4

[22] Dr. J Naga Padmaja, Amula Vijay Kanth, P Vamshidhar Reddy, B Abhinay Rao, Web Application for Emotion-Based Music Player using Streamlit, https://doi.org/10.22214/ijraset.2023.49019

[23] S Metilda Florence and M Uma, "Emotional Detection and Music Recommendation System based on User Facial Expression", IOP Conference Series: Materials Science and Engineering, vol. 912, no.6, pages:062007.