IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EMOTION BASED MUSIC RECOMMENDATION SYSTEM

[1]Renuka Mokalkar, [2]Uday Gaikwad, [3]Shubham Ghodake, [4]Hrutvik Jagtap, [5]Prof. Amol Jagtap

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor

[1]Computer Engineering,

[1]AISSMS COE, Pune, India

Abstract: The internet and mobile technology have developed quickly and made it possible for us to freely access various music resources. While the music industry might lean more toward certain genres of music. Our current playlists in the music listening apps are static and user explicitly have to change it based on their likings. Music recommendation systems have become a crucial part of the music listening experience. However, most traditional recommendation systems rely on user behavior data or metadata, which fail to capture the emotional content of music. As music has a strong emotional impact on listeners, personalized music recommendation systems that take into account the emotional state of users are highly desired. For this purpose there were two independent models and existing systems, one for detecting the mood from the facial expression that is FER(Facial Emotion Recognition) and another was Music Classification Models which were used to recommend a song. We are attempting to combine these two systems to make efficient and accurate recommendation to the user. So in this project, we are going to develop a system which will capture the real time emotion of user by conversating with user or by other means and based on that emotion related songs will be recommended. We are going to categorize songs into the groups based on the categories like Happy, Sad, Energetic and calm. Then according to the captured emotion from the user, the songs related to that emotion will be recommended. Key factor of the song recommendation is classification of songs based on their acoustic features. In this way, user can listen the songs according to the mood.

Index Terms - **Recommendation System, Facial Emotion Recognition, Interactive UI, Mood based music classifier.**

I. INTRODUCTION

1.1 Problem Statement:

Despite the large availability of music streaming services, users still face challenges in discovering music that matches their mood and emotions. Traditional music recommendation systems often rely on user behavior data or metadata, but fail to capture the emotional content of music. As music has a strong emotional impact on listeners, there is a need for personalized music recommendation systems that will give the solution to these factors. Sometimes the classification of songs is not up to the point which impacts the overall user experience. Therefore, this paper aims to develop a real time system that detects user's current emotions through facial expression and based on which it recommends music classified according to its acoustic features.

1.2 Purpose:

The purpose of an emotion-based music recommendation system is to enhance the music listening experience for users by providing personalized music recommendations that match their emotional state. Traditional music recommendation systems often rely on user behavior data or metadata and makes the overall rigid experience and fail to capture the emotional content of music.

Lyrical classification of songs may face several limitations in terms of language and context. The emotion-based music recommendation system aims to accurately identify the acoustic characteristics of music and match them with the user's emotional state, in order to recommend music that is appropriate for the user's mood. This can lead to a more enjoyable and satisfying music listening experience for users, as they are able to easily discover and listen to music that matches their current emotional state. Additionally, such a system could also help to introduce users to new music that they might not have otherwise discovered, based on the emotional content of the music rather than just the artist or genre and this will improve listening habits of user.

1.3 Scope:

By considering flexibility and feasibility of this project, it has capability to work as extended feature in the existing music listening apps as well as can be implemented as a whole independent system which will be based on user's emotion by improving certain aspects. In terms of the technical scope, the system involves the development of machine learning models for accurately identifying the emotional characteristics of music, as well as the integration of real-time emotion detection mechanisms. Additionally, the system would need to be scalable and adaptable to different music streaming services and platforms by ensuring additional user engagement can lead to the increment of the revenue.

II. LITERATURE REVIEW

1. Music Recommendation Based on Face Emotion Recognition : A person's emotions can be detected by the proposed system, and if the individual is feeling down, a playlist of the most upbeat, musically-related songs will be played. Additionally, if the emotion is positive, a specific playlist of songs from different musical genres will be made available to reinforce the good feelings.

2. Music Recommendation System using Content and Collaborative Filtering Methods : The primary goal of this proposed system is to increase the functionality of the current recommendation system. Utilizing collaborative filtering or content-based filtering, traditional music recommendation systems provide recommendations. Collaborative filtering and content-based filtering are combined in hybrid approaches to take advantage of both of their advantages and disadvantages.

3.Mood based Music Categorization System for Bollywood Music : The prime focus is to categorize the audio into different moods. Preprocess and feature Extraction will be done. After conducting survey and feature extraction process, Information Gain algorithm was used to select the defined features.

4.Chat bot song recommender system: Interaction with the system using chat bot that will assist user and recommend songs by analyzing emotion with the help of NLP methodologies.

5. Emotion Based Music Recommendation System: The majority of currently used music recommendation systems rely on content- or collaborative-based recommendation engines. However, a user's choice of music is not only based on past musical tastes or its actual substance. but also based on how that individual is feeling.

6. Music Recommender System for Users Based on Emotion Detection through Facial Features: we propose a recommender system for emotion recognition that is capable of detecting the user emotions and suggest a list of appropriate songs that can improve his mood.

7. Emotion recognition on FER-2013 face images using fine tuned VGG-16: Numerous studies have been done on this task, and each one has suggested a processing technique that is either standalone or ensemble-based. While many studies aim to improve accuracy, this study uses a standalone-based modified CNN (Convolutional Neural Network) based on VGG-16 (Visual Geometry Group 16) in an effort to improve efficiency.

8. Music recommendation system based on facial emotion recognition: This research focuses on creating an effective music recommendation system that analyses user facial expressions to identify their emotions. On a wider scale, this would result in the time and labor used to complete the operation manually being recovered. The system's overarching goal is to effectively identify facial expression and make music recommendations.

9. Music mood classification: This essay makes an effort to categorize songs according to the various emotional states they might be in, such as joyful, sad, furious, brooding, quiet, uplifting, etc. A fascinating and difficult subject with several applications in the music business is the creation of a framework for the estimate of musical mood that is resilient to the extreme diversity of musical material across genres, performers, global areas, and historical periods.

III. METHODOLOGY

3.1 System Architecture

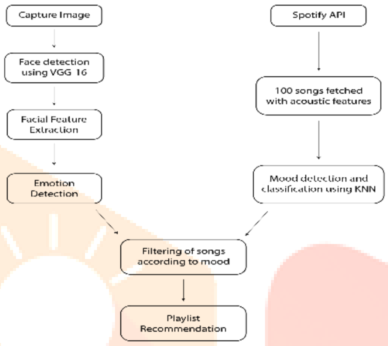


Figure 1: System Architecture

3.2 Data Gathering:

Data gathering is a critical step in the machine learning process, as it provides the raw material for training, testing, and validating machine learning models. The quality and quantity of data gathered plays a crucial role in determining the accuracy, robustness, and generalization ability of machine learning models. Efficiency and accuracy of the project depends on the factors of the dataset such as number of features, complexity and diversity of data. The more diverse the dataset more effective outcomes will be.
In summary, data gathering is important in machine learning as it directly affects the accuracy, robustness, and generalization ability of the machine learning models. Gathering diverse and representative data is crucial for ensuring that the models can generalize well to new data and can be used to solve real-world problems.

As our system attempts to combine two models which are FER(Face Emotion Recognition) and Music Classification.

Dataset used for Facial Emotion Recognition model –
There are some datasets present for the FER model like FER-2013, CK+, AffectNet JAFFE and EmoReact. As we can see CK+,JAFFE and EmoReact has very less number of images which can affect testing accuracy and in case of AffectNet, it has crowdsourced annotations may contain noise and biases, potential issues with imbalanced classes, limited diversity in age, gender, and ethnicity. So we finalized FER-2013 dataset for our FER model.

Table 1: Dataset Comparison

Dataset	Number of images	Number of emotions	Image size	Annotation method	Year
FER 2013	35,887	7	48×48	Crowdsourcing	2013
CK+	981	6	Various	Expert annotation	2000
JAFFE	213	7	256×256	Expert annotation	1997
EmoReact	1,840	7	Various	Crowdsourcing	2017
AffectNet	1,000,000	8	Various	Crowdsourcing	2017

FER2013: The dataset contains 35,887 grayscale images of size 48x48 pixels. Each image is labeled with one of seven emotion categories: angry, disgust, fear, happy, sad, surprise, or neutral. The labels were obtained using crowd-sourcing, where human annotators were asked to classify the emotions expressed in the images. The dataset is divided into a training set with 28,709 images and a test set with 3,589 images.

Dataset used for Music Classification Model-
Music Mood Dataset: While researching about our project, we came across various datasets that can be used for the classification purpose. But, we didn't find the factors which we were looking for our project. Later we found one dataset that was created by certain person for his research work because the songs were classified into multiple acoustic features such as Loudness, Instrumentelness, Danceability, Acousticness, Valence etc. and is labelled by well defined categories of songs. The labels of this dataset are totally aligned with our categories of mood. So we decided to go with this dataset. This dataset contains sufficient amount of data for each category so the prediction will be accurate and efficient. This dataset contains 687 rows as they denote multiple songs. It contains around 19 features in which certain features elaborate more about information related to the song such as Name, Album, Artist, Popularity etc. and the rest are all its acoustic features.

3.3 Model Selection:

Our project predominantly consists of two major models.
One is for FER(Facial Emotion Recognition) and another one is for Music Classification.
For FER model, we researched about various techniques and methodologies that can be followed. The obvious model or the Image Recognition in the deep learning is CNN(Convolutional Neural Network).

CNN:
The basic building block of a CNN is a convolutional layer, which applies a set of filters to an input image and produces a set of output feature maps. These filters learn to detect specific patterns or features in the input image, such as edges, corners, or textures. By stacking multiple convolutional layers, a CNN can learn more complex features and patterns at different levels of abstraction.

But it might give inaccurate predictions as the complexity of the model gets increased. So we decided to go for the pretrained models for FER by taking the factor of accuracy into consideration. Some of the popular pretrained models out there mainly includes ResNet-50, VGG-16, Inception-V3, MobileNet-V2 etc. Major disadvantage of MobileNet-V2 is that it may not capture complex and abstract features as well as larger models. In case of Inception-V3 and ResNet-50 both may require high computational resources for training and inference and may be prone to overfitting. As the main requirement of our project is to do the complex computations and accurate prediction which best fits for the VGG-16 model and not in the case of other pretrained models. In addition to this, theoretically we found that VGG-16 has highest accuracy of 71% among other pretrained models which have accuracy in the range of 68% to 70% on FER-2013 dataset. Hence we headed with the VGG-16 model.

VGG-16:

In FER (Facial Emotion Recognition), VGG-16 refers to a specific architecture of Convolutional Neural Network (CNN) that is used for facial expression recognition from images or videos. VGG-16 has a deep architecture with 16 layers, allowing it to capture more complex and abstract features from the input image. This is particularly useful for FER, as emotions can be subtle and difficult to detect.
It has been pre-trained on the large-scale ImageNet dataset, which contains millions of labeled images across 1000 classes. This pre-training allows the model to learn rich feature representations that can be fine-tuned for specific tasks such as FER. By fine-tuning the pre-trained VGG-16 model on a FER dataset, the model can learn to extract relevant facial expression features and classify them into different emotion categories such as happy, sad, angry, or surprised.

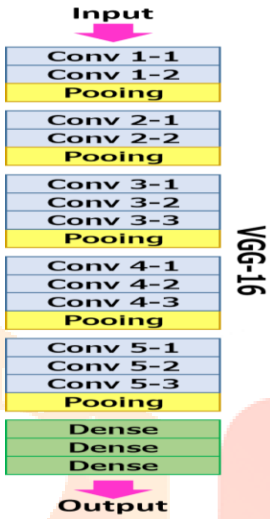


Figure 2: VGG-16 architecture

In case of our project, To train lesser parameters at certain instance VGG-16 uses convolution layer kernel of size 3x3 instead of 5x5 or 7x7 kernel. Moreover it uses two or three stacked convolution layers with 3x3 kernel to make similar for the one convolution layer with 5x5 or 7x7 kernel respectively. A discriminative decision function is resulted by adding more rectification layers which predicts output with more accuracy. It has done well, with several of the tested datasets—including the FER-2013 dataset—achieving SOTA (State of the Art) results.

For Music Classification model, the main dataset upon which we are relying is Music Mood Dataset as we discussed about it in previous sections. It is a well labeled dataset having four main classes. Considering these factors of our projects some of the prominent classification algorithms are KNN(K-Nearest Neighbors), SVM(Support Vector Machine), Naïve Bayes, Decision Trees and Neural Networks etc. By considering pros and cons of all above mentioned algorithms, we decided to choose KNN algorithm based on its simplicity and effectiveness. Thus it is the perfect choice for our music recommendation system.

KNN(K-Nearest Neighbors):

KNN algorithm is based on similarity measures between the attributes of two instances. In the context of music recommendation, this algorithm can be used to find songs that are similar to the ones that the user likes. KNN is a non-parametric algorithm, which means it does not make any assumptions about the underlying distribution of the data. It can handle complex decision boundaries and is suitable for datasets with a large number of features. However, it is computationally expensive during prediction as it requires calculating distances to all the training instances.

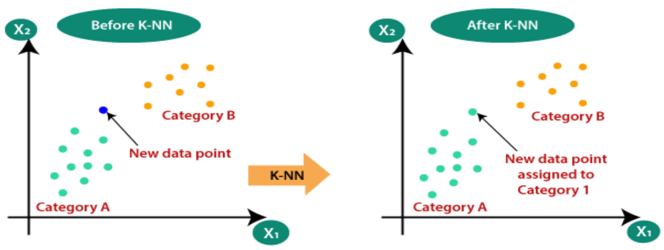


Figure 3: Working of KNN

Why KNN?

Digging down about the various algorithms, we came across a thing that Neural Networks and SVM are computationally expensive and may take extra resources. Despite of its efficiency and ability to work with the multidimensional data, Naïve Bayes is not capable to capture the complex relationships between the features. Sensitivity to the small variations and the problem of overfitting can be there with the Decision Trees which will not fit in the frame of requirements of our project. Hence taking all these factors in consideration, decision of taking forward KNN model for music classification model is taken.

Table 2: Algorithm accuracy comparison

Algorithm	Accuracy
KNN	High
SVM	High
Random Forest	High
Naive Bayes	Medium
Logistic Regression	Medium
Decision Trees	Low

Thus, here we finalized our models by which we will be proceeding into the next phase.

3.4 Model Building:

Model building is a crucial step in the machine learning process as actual development and implementation of theoretically proposed models and methodologies are implemented in this phase. The model we proposed or decided in the previous phases is trained on the respective datasets and accordingly it is used to give prediction on the new data.

Here as we discussed earlier we are going to work on FER and Music Classification Model. But the prerequisite step in any machine learning process before actual development of model is Data Preprocessing where certain operations are performed on the datasets to make it ready for model training.

Data preprocessing is an essential step in machine learning that involves transforming raw data into a format that can be easily analyzed and used by machine learning algorithms. It involves a series of steps that aim to clean, normalize, and transform data before it is used for training and testing machine learning models. As we proceeded with FER model, there were requirements of our project to convert the data into certain format to make it easily analyzed and this is nothing but the Data Transformation. The obvious input of the FER model is image which existed in the dataset in various formats such as .PNG and other. so we converted those image formats into the form of array to make it easy to analyze for the model. Another technique used for data preprocessing on FER-2013 dataset which involves scaling the data so that it falls within a specific range or distribution. Here we converted values into the range of 0-1. So this can help us to improve the performance of our algorithm. Major techniques of data preprocessing on other model which is Music Classification are Data cleaning and Feature selection which includes removing or correcting any errors, inconsistencies, or missing values in the data and selecting the most relevant features or variables that are likely to have the most predictive power for a given machine learning problem respectively.

FER Model:

When user gives his/her image as a input to the system, it is processed further frame by frame. As we finalized VGG-16 model for the facial emotion recognition, we are using modified VGG-16 architecture with GAP(Global Average Pooling) as a final pooling layer. If we take 512 averaged input neurons then we can replace all the classifier layers of VGG-16 by a single fully connected layer and accordingly we can get 7 neurons as the similar to that of classes of FER-2013 dataset which mainly consists moods categorized into 7 different categories.

Before going to perform actual training on the model, the loading of required datasets and libraries is done. Later we analyzed, the actual mood configuration that is how much number of images are present for the respective mood category. After analysis we come to know that images of different moods are present in different proportion. In case of FER-2013 dataset, the images it contains have the size as 48x48, so we decided to resize our input image in the similar size and this will eventually help to get accurate prediction.

For training purpose we splitted our dataset into 80% as training and remaining 20% as testing data ensuring this splitting is totally random. Transfer learning is a technique in deep learning where a pre-trained model is used as a starting point for a new task. Instead of building and training a model from scratch, transfer learning allows us to leverage the knowledge and learned representations of an existing model. VGG-16 can be used for transfer learning, where the pre-trained model is fine-tuned on a smaller dataset for a specific task such as FER. Transfer learning is a popular approach in deep learning that allows for faster and more efficient training of models on new datasets. We used Softmax activation function which is often used in the output layer of a neural network for multiclass classification tasks. It transforms a vector of real numbers into a probability distribution over the classes.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 1, 1, 512)         14714688

 global_average_pooling2d (G  (None, 512)              0
 lobalAveragePooling2D)

 dropout (Dropout)           (None, 512)               0

 dense (Dense)               (None, 7)                 3591

=================================================================
Total params: 14,718,279
Trainable params: 14,718,279
Non-trainable params: 0
_________________________________________________________________
```

Figure 4: Model Architecture

The above snapshot depicts the architecture of VGG-16 along with its shape and parameters. Global Average Pooling (GAP) is a type of pooling layer in convolutional neural networks (CNNs) that takes the average value of each feature map in the previous layer to generate a single value for each feature map. This pooling operation reduces the dimensionality of the feature maps, making them more compact and easier to process. Using this faster training and better generalization is ensured. A regularization method used to lessen overfitting of the model is the dropout layer. It is a type of regularization technique that is used to prevent neural networks from overfitting the data. The dropout layer works by randomly dropping out a certain percentage of the neurons during each training iteration. During training, each neuron in the dropout layer has a probability p of being dropped out. For our model, the probability p is set to 0.5. A Dense layer is a neural network layer that is used for feedforward networks in deep learning. The dense layer performs a linear operation on the input followed by a non-linear activation function, such as a ReLU or sigmoid function. It computes a weighted sum of the input features along with a bias term and applies an activation function to generate the output of the layer. After performing the training operation on the model it can give the accurate output.

Music Classification Model:

Base of music classification model is the music mood dataset which we have used. It contains several categories of songs that are ultimately mapped to the input emotion. This model is trained on this dataset and we made it ready to make the prediction on the new data. Model training starts with importing all the necessary libraries such as Numpy, Keras, Seaborn and Matplotlib etc. followed by loading dataset. The four main labels of our dataset are into text format so we encoded them into the integers from 0 to 3 to make it easy for mapping to the mood which we got as a input from user. Further the splitting of our dataset is done into two parts that is training and testing. Here we divided 80% of our data for the training purpose and remaining 20% data for testing purpose. This splitting typically done for avoiding the overfitting problem where model fails to fit the data. Hence we trained our model using KNN algorithm and our model is ready to make prediction.

Now the key factor of our dataset is that it contains songs with there respective acoustic features. Instead of using static dataset or constant data to the model and making predictions on these types of data can affect the flexibility of the system. So we started looking to the ways where we can provide the real time and dynamic song data to the model, and then trained model using KNN will make predictions on this data resulting into overall satisfying user experience and variety of songs will be recommended. For fetching real time songs, instead of using large dataset which contains thousands of songs, we decided to use Spotify API and Spotipy library to generate those real time songs.

Spotify API provides some of the additional functionalities to the developers which can help them to develop excellent projects. These functionalities mainly includes retrieval of data from targeted album, song or artist, searching the desired content, managing personal libraries and control and interaction. In addition to this it can provide us variety of songs along with audio features of each track. This audio features are nothing but acoustic features which are also features of our dataset. Spotipy, official library of Spotify deals with these track features.

Why Acoustic analysis over lyrical?

Several studies and researches are based on lyrical classification in which they attempts to find exact emotion of the text based on which classification of that song is done. We found that emotional interpretation of lyrics can vary from person to person and exact emotion can be different from interpreted emotion. Moreover, this type of analysis makes the scope of recommendation limited to language. By these finding we decided to continue with Acoustic feature analysis as they can provide more objective and reliable classification as it is based on measurable acoustic features and can work with instrumental music, where lyrics are not present. Thus making the scope of the system more global irrespective of the language barriers as each and every song has certain magnitude of acoustic features with varied impact.

This concept we molded into our project in such manner that we fetch 100 songs through this Spotify API whose track features are matching with the features of our dataset. Each feature has certain magnitude which denotes value and has influence on the overall mood of the song. Model trained with KNN algorithm classifies this newly generated song data into predefined 4 categories and those are Happy, Sad, Calm, Energetic. Thus the expected category of songs is recommended based on the input mood of the user.

```
labels = ['Calm', 'Energitic', 'Happy', 'Sad']
# emotion_map = ['Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprise', 'Neutral']
mood_to_song_map = [0,0,0,2,3,1,0]
```

Figure 5: Mood Mapping 1

As given in FER model we get an input integer which is nothing but the mood and based on the mapping of the detected emotion and mood of the song the playlist of the mood will be recommended.

The mapping here we do in is such manner if user is happy then he/she will get recommended happy songs, if user is sad he/she will get recommended sad songs. Energetic songs will be recommended if user is in surprise mood. For angry, disgust, fear and neutral mood the calm songs will be recommended.

Figure 6: Mood Mapping 2

IV. MODEL EVALUATION

Model evaluation is the process of assessing the performance of a machine learning model. It involves comparing the predicted outputs of the model with the actual outputs and determining how well the model has performed. The aim of model evaluation is to determine the accuracy and reliability of the model, identify any issues or areas of improvement, and optimize the model for better performance.

There are several metrics that can be used to evaluate the performance of a machine learning model, depending on the type of problem and the type of model. In case of our model, we are evaluating our model with the help of performance metrics like accuracy, precision, recall, F1 score etc. These metrics provide a quantitative measure of the model's performance and can be used to compare different models or variations of the same model. Performance of our model gauged by these evaluation metrics is given below.

	precision	recall	f1-score	support
0	0.60	0.62	0.61	991
1	0.68	0.62	0.65	109
2	0.58	0.47	0.52	1024
3	0.89	0.88	0.88	1798
4	0.56	0.57	0.57	1216
5	0.77	0.77	0.77	800
6	0.62	0.71	0.66	1240
accuracy			0.69	7178
macro avg	0.67	0.66	0.67	7178
weighted avg	0.69	0.69	0.68	7178

Figure 7: Model Performance

The accuracy we achieved with FER model is 69%. Macro average for precision, recall and F1 score lies around 67% and weighted average for the same lies around 69%.

We performed visualization and plotted graph for model accuracy and model loss which is used to provide an intuitive understanding of the model's performance.

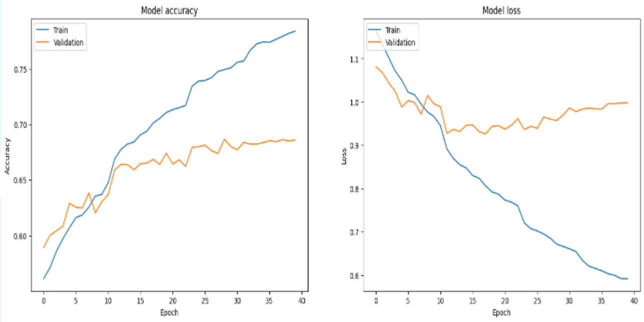


Figure 8: Model Accuracy and Loss Plot

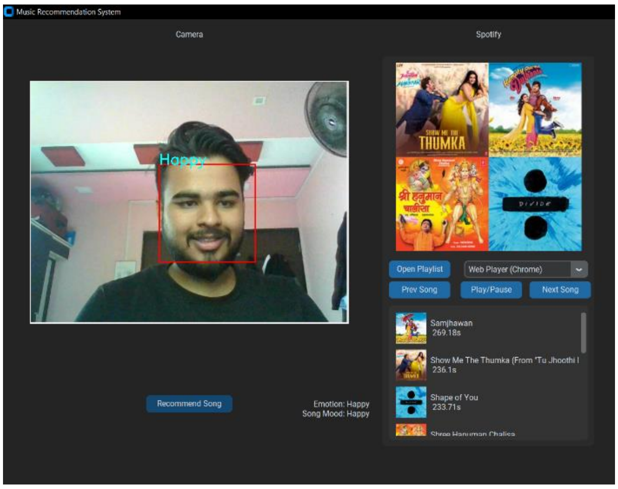
Here we can observe that, as number of epoch increases the accuracy also increases. We got the best accuracy at epoch number 27 and there after it started decreasing. In case of the loss as number of epoch increased loss decreased till epoch number 18 after that it stared increasing. So we recorded both these results for our model.

In case of music classification model, we plotted confusion matrix to analyze performance of KNN. It is a table used to evaluate the performance of a classification model by comparing the predicted and actual labels. It give us idea about the predicted values and true values of the 4 labels named as Happy, Sad, Energetic and Calm. By taking these values in account we can evaluate the performance of the model. Thus we got 79% model accuracy.

Accuracy Score 0.7898550724637681

Figure 9: Confusion Matrix

V. RESULTS



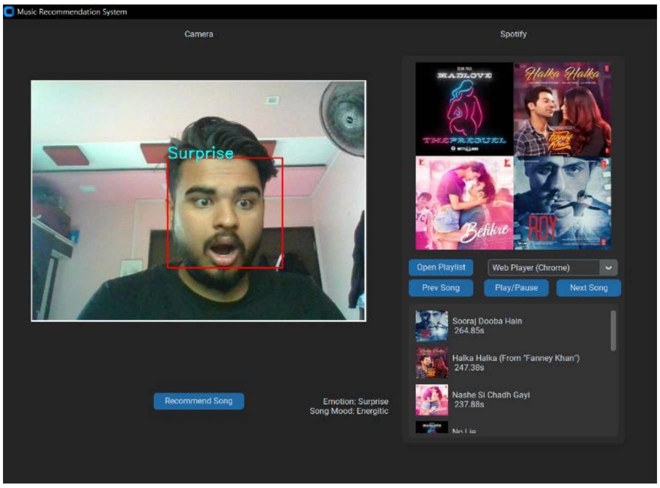
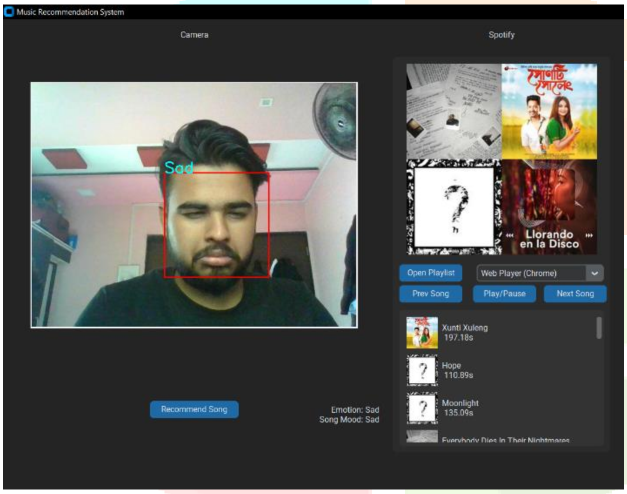


Figure 10: Recommendation for Happy Mood Figure 11: Recommendation for Surprise
Mood



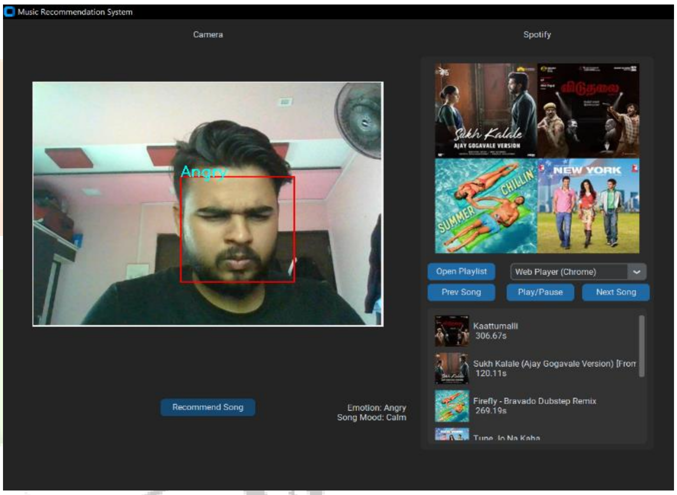


Figure 12: Recommendation for Sad Mood Figure 13: Recommendation for Angry
Mood

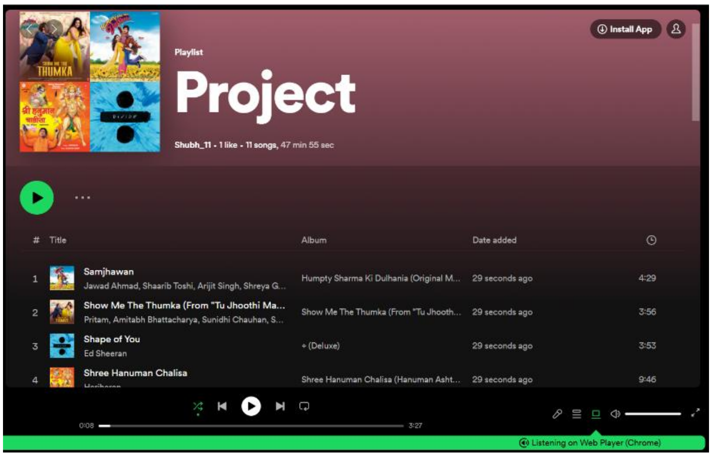


Figure 14: Spotify Playlist Recommendation

VI. FUTURE SCOPE

This system is fully operational, however there is still opportunity for advancement. Numerous changes can be made to the application to enhance user experience overall and produce better outcomes.

Currently existing systems face difficulties while giving manual inputs but our system performs well when there is difficulty in giving manual inputs as we are dynamically determining the mood of a person from its expressions. In addition to this, this system not only used as a feature but also as a independent system. If further technical advancements done and if it improved the efficiency of the algorithm used then we can also recommend the songs on crucial and trivial expressions. [6] Current system mainly focuses on doing the recommendation on content based features that is based on its acoustic parameters. But in future hybrid model can be created that uses both content based and collaborative(user history, feedback, liking etc.) for recommendations. Accuracy of the face recognition model can be further improved by using more complex computations and using more deep layers of the network.

The system's future plans include designing a mechanism that could aid in the treatment of individuals who are experiencing mental stress, anxiety, acute depression, and trauma through music therapy. Due to the current system's poor camera resolution and performance in extremely low light levels, there is a chance to add some functionality as a potential fix in the future. If the camera has poor lighting, we can either offer music based on the user's general facial expressions or, as we already indicated, we can employ collaborative features to suggest the right song to the user. As this problem is more inclined towards hardware issue we can overcome it by using high quality camera devices and sensors having high end technical specifications.

VII. CONCLUSION

A detailed analysis of the literature reveals that there are various ways to put the Music Recommendation System into practice. The methods proposed by preceding researchers and developers were examined. So when we started studying we mainly found 2 approaches and that too independent. The first approach was like just determining the accurate emotion from the facial expression and second one was classifying the songs into the front emotions based on their acoustic features. Based on the outcomes, we fixed the system's objectives. So we decided to merge these two approaches and provide complete solution for existing problem. The system in place has the ability to detect user emotions. The system was able to identify happy, sad, angry, neutral, disgust, fear and surprise emotions. The suggested method after identifying the user's emotion provided the user with a playlist of music matches that matched the user's emotion. Memory and CPU usage increase as a result of processing a large dataset. Development will become more difficult and appealing as a result. The goal is to develop this application as affordably as feasible and on a common platform. Users will find it simpler to create and manage playlists with the help of our facial emotion-based music recommendation system.

REFERENCES

[1] AYUSH Guidel, Birat Sapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020.

[2] Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Re-search & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[3] Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, an intelligent music player based on emo-tion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

[4] CH. sadhvika, Gutta.Abigna, P. Srinivas reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is-sue 4, April 2020.

[5] Madhuri Athvale , Deepali Mudale, Upasana Shrivatsav, Megha Gupta, Music Recommendation based on Face Emotion Recognition, Department of Computer Engineering, NHITM, Thane, India

[6] Sheela Kathavate, Music Recommendation System using Content and Collaborative Filtering Methods, Department of Information Science and Engineering BMS Institute of Technology and Management Bangalore, India

[7] Amey Ujlambkar, Omkar Upadhye, Akshay Deshpande, Gunwant Suryawanshi, Mood Based Music Categorization System for Bollywood Music.

[8]Jose Padial. Ashish Goel, Music Mood Classification.

[9] Emotion Recognition on FER-2013 Face Images Using Fine-Tuned VGG-16. Gede Putra Kusuma, Jonathan, Andreas Pangestu Lim.

[10] Deep-Emotion: Facial Expression Recognition
Using Attentional Convolutional Network. Shervin Minaee, Amirali Abdolrashidi

[11] Challenges in Representation Learning: A report
on three machine learning contests.

[12] Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity Tanoy Debnath, Md. Mahfuz Reza, Anichur Rahman, Amin Beheshti, Shahab S. Band & Hamid Alinejad-Rokny