MUSIC RECOMMEDER SYSTEM BASED ON FACE EMOTION DETECTION

A PROJECT REPORT

Submitted by

SHREYANSH AGRAWAL (20BAI10084)

NIKHIL CHAURASIYA (20BAI10106)

ANJALI MAHESHWARI (20BAI10198)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING WITH SPECIALIZATION IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

VIT®
BHOPAL
www.vitbhopal.ac.in

SCHOOL OF COMPUTING SCIENCE AND ENGINEERING

VIT BHOPAL UNIVERSITY

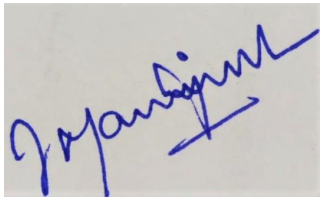
KOTHRIKALAN, SEHORE

MADHYA PRADESH - 466114

APRIL 2022

BONAFIDE CERTIFICATE

Certified that this project report titled **"MUSIC RECOMMENDATION SYSTEM BASED ON FACE EMOTION RECOGNITION"** is the bonafide work of "**SHREYANSH AGRAWAL (20BAI10084), NIKHIL CHAURASIYA (20BAI10106), ANJALI MAHESHWARI (20BAI10198)"** who carried out the project work under my supervision. Certified further that to the best of my knowledge the work reported at this time does not form part of any other project/research work based on which a degree or award was conferred on an earlier occasion on this or any other candidate.



U.G. PROGRAM CHAIR
Dr. S SOUNTHARRAJAN
Associate Professor
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

PROJECT GUIDE
Dr. MANIKANDAN J.
Assistant Professor
School of Computer Science and Engineering
VIT BHOPAL UNIVERSITY

ACKNOWLEDGEMENT

First and foremost, I would like to thank the Lord Almighty for His presence and immense blessings throughout the project work.

I wish to express my heartfelt gratitude to **Dr. S SOUNTHARRAJAN**, Head of the Department, School of Computer Science and Engineering for much of his valuable support encouragement in carrying out this work.

I would like to thank my internal guide **Dr. MANIKANDAN J.**, for continually guiding and actively participating in my project, giving valuable suggestions to complete the project work.

I would like to thank all the technical and teaching staff of the School of Aeronautical Science, who extended directly or indirectly all support.

Last, but not least, I am deeply indebted to my parents who have been the greatest support while I worked day and night for the project to make it a success.

LIST OF ABBREVIATIONS

ABBREVIATIONS	MEANING
EDA	Exploratory Data Analysis
RNN	Recurrent Neural Network
PLSA	Probabilistic Latent Semantic Analysis
ANN	Artificial Neural Network
CNN	Convolution Neural Network

LIST OF FIGURES AND GRAPHS

ABSTRACT

Face recognition technology has gotten a lot of press because of its vast range of applications and business possibilities. It is used in a variety of sectors, including security systems, digital video processing, and other technical advancements. Furthermore, music is a type of art that is seen to have a stronger emotional connection. It has a one-of-a-kind power to improve one's mood. In general, this project focuses on developing an effective music recommendation system that leverages Facial Recognition algorithms to assess the user's sentiment.

The constructed method would prove to be more effective than previous systems. Furthermore, on a wider scale, this would allow for the recovery of time and labour spent physically conducting the procedure. The system's overall goal is to identify facial expression and quickly propose tunes. Both time and money will be saved with the suggested approach.

Our music player contains three modules: *Emotion Module, Music Classification Module and Recommendation Module.*

The Emotion Module uses deep learning algorithms to assess the user's mood based on a picture of their face as input. The Music Classification Module employs audio elements to obtain a surprising outcome when categorizing music into four distinct mood groups.

The Recommendation Module recommends music to the user by mapping their feelings to the song's mood type and taking into account the user's preferences. With the use of a camera, the user's image is taken. The user's photo is captured, and then a suitable music from the user's playlist is presented that matches the user's requirements, based on the user's mood/emotion.

TABLE OF CONTENTS

INTRODUCTION

People's emotions are mostly expressed through their facial expressions. Music has always been recognised to change a person's mood. Capturing and identifying a person's emotion and showing appropriate tunes that match the person's mood can help to soothe the user's mind and provide a pleasing overall impression. The goal of the project is to capture a person's feelings through their facial expressions.

Through the web camera interface available on computing systems, a music player is meant to record human expression. The software captures the user's image and then extracts features from the face of a target human being using image segmentation and image processing techniques in order to discern the emotion that the person is attempting to express.

The project seeks to brighten the user's mood by playing tunes that match the user's criteria by collecting the user's photograph. Facial expression recognition has been the best type of expression analysis known to humanity since ancient times. People's facial expressions are the most effective technique of analysing or concluding the emotion, feeling, or thoughts that another person is attempting to express. Mood change may also aid in the treatment of depression and melancholy in some cases. Many health concerns can be prevented with the use of expression analysis, and efforts can also be made to improve a user's mood.

MOTIVATION

We must first understand how important this problem is to real world scenarios. Let's see few applications where a solution to this problem can be very useful:

Feeling and character are essential mental constructs.Emotions are momentary emotional reactions to a particular upgrades, yet personality qualities are an anticipated and steady measure that decides human behaviour. Both an affect music inclinations and MRSuser requirements. Character and feeling, then again, have yet to play a significant part in (music) recommender frameworks. We accept that psychologically inspired MRS is an impending region.

Self-Career Advancement - Feeling acknowledgment has gained a lot of importance in all parts of life and on the off chance that a powerful calculation implemented which can accurately arrange the feelings of the individual, a lot of progression in the industry can be accomplished with the assistance of this. The framework has successfully been able to catch the feeling of a client. It has been tried in a continuous climate for this predicate. In any case, it must be tried in various lighting conditions to determine the robustness of the created framework. The framework has additionally had the option to grab the new images of the client and suitably update its classifier and preparing dataset. The system was planned utilizing the facial tourist spots plot and was tried under various scenarios for the outcome that sounds acquired, really. It is seen that the classifier has an accuracy of in excess of 80% for the greater part of the experiments, which is pretty good accuracy regarding feeling grouping. It can likewise be seen that the classifier can accurately anticipate the outflow of the client in a constant situation when tested live for a client.

OBJECTIVE / PROBLEM STATEMENT

The goal of Music Recommender Systems is to recommend songs to the target user in a personalized manner. Song Recommendation is a challenging artificial intelligence problem where an analytical approach must be employed for the given dataset.

The task of music recommendation can be divided into two modules logically.

1. Pre-processing
2. Training (creating the final model)

The main objective of this work is to develop an application that reads our facial expressions and recommend music according to our mood. Whenever a user looks in the camera while using the app, it reads their expressions and suggest songs accordingly. In order to do that we use various strategies to implement recommendation engine.

LITERATURE REVIEW

S.No.	Title of the paper	Journal Name, Year of Publication and volume & issue number	Author Name	Problem Addressed / Problem Statement	Methods / Technologies used	Author Contribution	Shortcomings / Deficiency / Assumption made
1.	Emotional Detection and Music Recommendation System based on User Facial Expression	Emotional Detection and Music Recommendation System based on User Facial Expression	S. Metilda Florence and M. Uma	Recommend songs on basis of User's Facial Expression.	• Emotion Extraction Module • Audio Extraction Module • Emotion - Audio Integration Module	Author's contribution was he made a scenario where predictions were performed on emotions expressed	The image that is fed into the classifier should be taken in a well-lit atmosphere for the classifier to give accurate results.
2.	Music recommendation system based on facial emotion recognition	Music recommendation system based on facial emotion recognition.	Deny John Samuvel, B. Perumal, Muthukumaran Elangovan	The overall concept of the system is to recognize facial emotion and recommend songs efficiently.	A system which will recommend music by recognizing the mood of the user from facial emotions is the overall concept described in the paper.	When multiple faces are considered, they are compared by detecting these parts of the faces	The quality of the image should be at least higher than 320p for the classifier to predict the emotion of the user accurately.

RELATED WORK

To successfully group human feelings, a number approaches have been presented and adopted. The majority of the techniques focused on seven basic sensations that are consistent across age, culture, and distinct characteristics. This paper explains the advantages of utilising OpenCV, particularly the Adaboost algorithm, in the face recognition process. A combination of a specific algorithm and the AdaBoost algorithm can be used to detect and recognise faces in complex colour photographs. It also discusses the drawbacks of utilising a timer for facial detection.

To arrange eight facial feelings, the author proposes using Support Vector Machines (SVM) as the major characterization technique. The faces were identified using channels in OpenCV and then converted to gray scale. The research also discusses on robotized constant coding of external appearances in nonstop video gushing, which is applicable to apps that accept frontal viewpoints via web-cam.

The developer provided a formula for generating a subset of a custom playlist or playlist related to the emotion felt. The image to be created was taken with a webcam or directly from the hard circle. The image is subjected to enhancements, including a few mapping and upgrade methods to restore the image's required distinctiveness. To encourage multi-class characterization, the "oneversusall" strategy of SVM is used to preserve preparation and organisation.

Proposals for the use of deep convolutional neural networks. It is based on robust face recognition convolutional networks that can be effectively modified to perform the feeling recognition challenge. For enhanced face recognition, visual models are augmented by sound highlights. Contributes to the music suggestion framework, which is also a key component of the suggested framework. It addresses which elements of the music should be deleted in order to identify the music's mood. The paper shows how to interpret the state of mind of a musical piece using Thayer's model of mind-sets. The edge level of a piece of music is resolved, and the emotion it evokes is sensed through prepared brain networks.

REQUIREMENT ARTIFACTS

Hardware required...

Hard disk (SSD Preferred)

At least 16 GB RAM

Multicore GPU

min. processor i7

Software required...

- Python and Conda virtual environment
- Deep Learning concepts like:
  - CNN
  - RNN
  - ANN
- Spotipy Library
- Spotify API
- Scikit Learn library
- Tensorflow 2.5
- Numpy Library
- Pandas Library
- Seaborn Library
- Matplotlib Library
- Keras Library
- NLP knowlegde

METHODOLOGY AND ARCHITECTURE

Methodologies

- **10.1. Emotion Extraction Module** -The picture of the client is captured with the help of a camera/web-cam. When the image caught, the casing of the captured image from web-cam feed is changed over completely to a gray scale picture to improve the performance of the classifier, which is utilized to recognize the face present in the picture. When the transformation is finished, the picture is shipped off the classifier algorithm which, with the assistance of component extraction procedures can extract the face from the casing of the web camera feed. From the extricated face, individual features are acquired and are shipped off the prepared organization to recognize the emotion expressed by the client. These pictures will be utilized to prepare the classifier so that when a totally new and obscure arrangement of pictures is introduced to the classifier,it can remove the place of facial milestones from those pictures based on the information that it had proactively gained from the preparation set and return the coordinates of the new facial tourist spots that it identified. The network is trained with the assistance of CK broad informational collection. This is utilized to distinguish the emotion being voiced by the client.

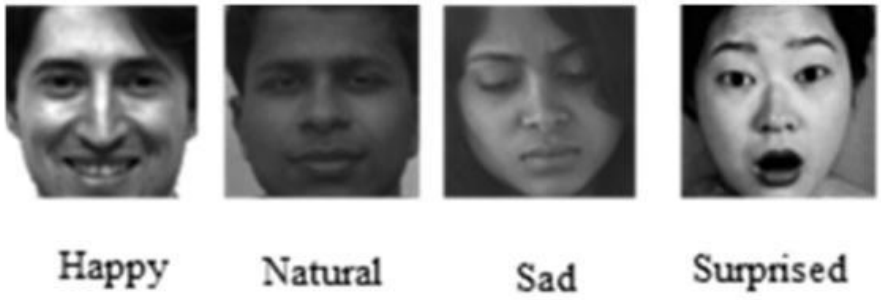


Fig. 1 Example Images on some emotions

- **10.2. Audio Extraction Module** - After the feeling of the client is extracted the music/sound in light of the inclination voiced by the client is shown to the user, a list of tunes in view of the inclination is shown, and the client can listen to any song he/she might want to. In light of the consistency that the client would listen to the melodies are shown in a specific order. This module is developed using web technologies like PHP, MySQL, HTML, CSS, JAVASCRIPT.

- **10.3. Emotion - Audio Integration Module** - The extracted emotions for the songs are saved, and the songs based on the emotion are displayed on a PHP and MySQL-based web page. For example, if the mood or facial characteristic is classed as happy, the user will see music from the cheerful collection.

Aimed Architecture:

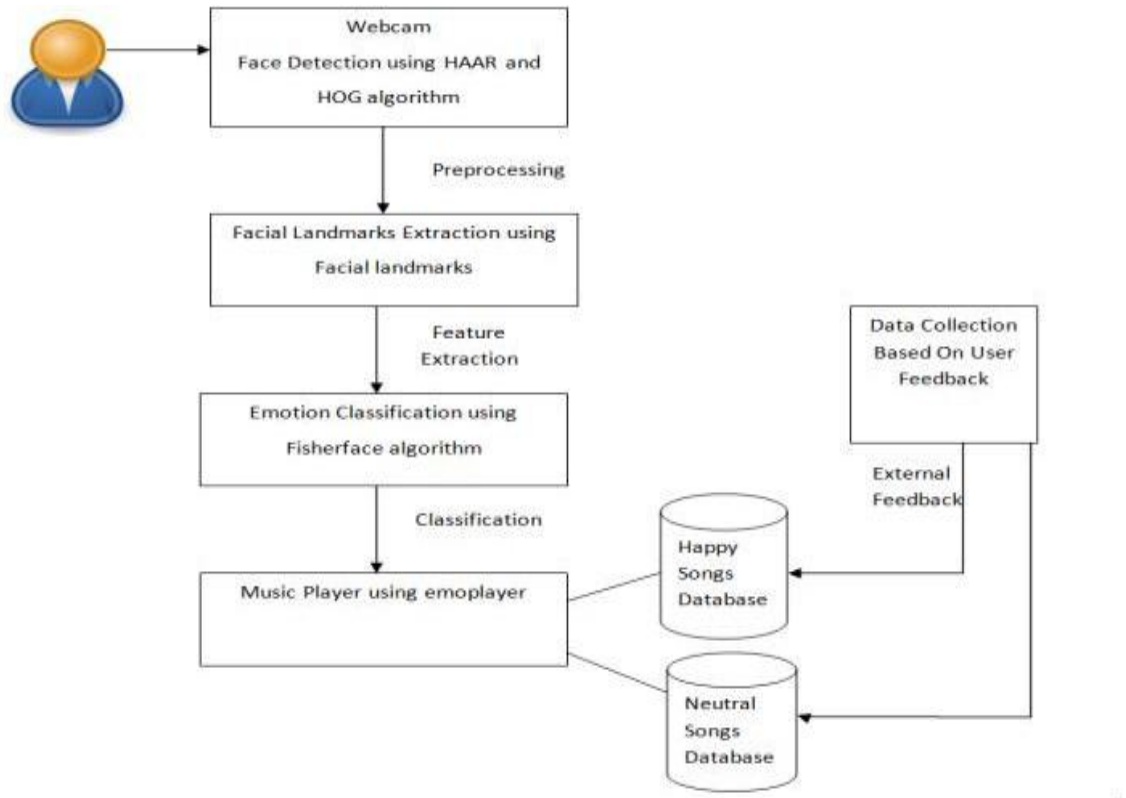


Fig 2 (The aim was that final Architecture to be looking like the above structure)

10.4. Data Understanding by Visualization and EDA

Visualization of the different songs data into year basis or the repetitive basis over a period of time Using the data grouped by year, we can understand how the overall sound of music has changed from 1921 to 2020

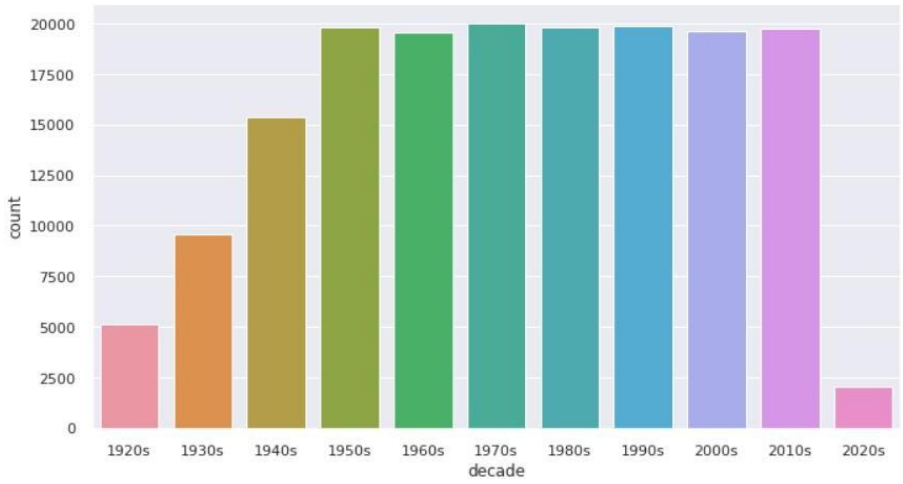


Fig. 3 Exploring Dataset of spotify which song was heard in which year

10.5. FlowChart

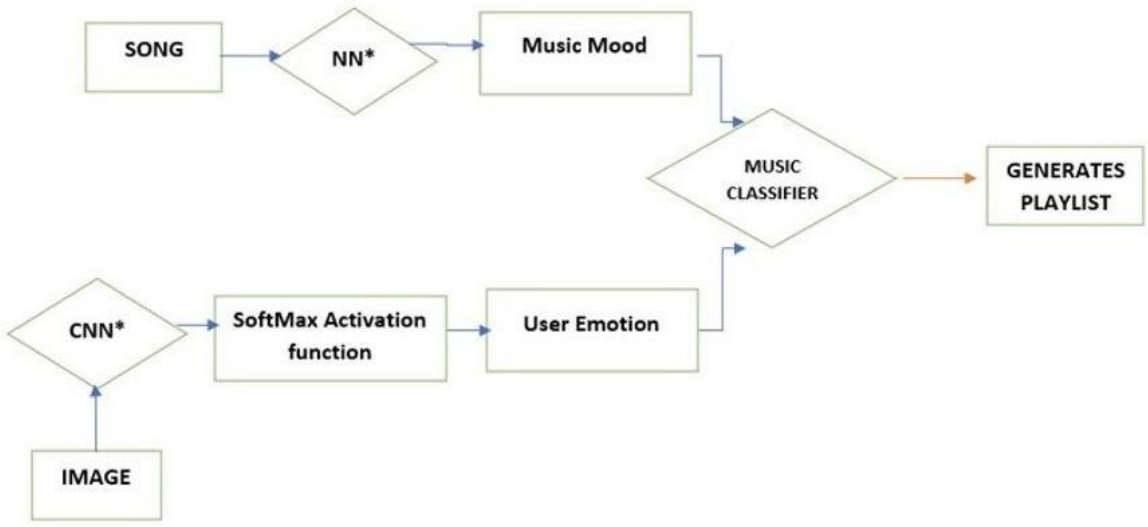


Fig. 4. Flowchart of our model

10.6. Clustering Genres with K-Means

To propose songs based on song metadata, we used the K-Means algorithm. We first built a space of songs based on several metadata aspects (artist, genre, etc.) and then recommended comparable songs. The Euclidean distance between each pair of vectors is then determined, and the vectors are then clustered into groups using the K-Means technique. Finally, each group's profile is created by aggregating the song, artist, and genre entries made by that group's users, then averaging and rounding the user's ratings for loudness, energy, happiness, tempo, popularity, and rhythmic prominence listed by that group's users.

10.7. Spotify Playlist Recommendation:

The proposed system will display the appropriate music playlist based on the identified emotion. The classical music playlist will be active for pleasant emotions, while the new age music playlist will be activated for natural emotions. For unpleasant emotions such as surprise and sadness, a designer music playlist will be enabled to improve the user's mood. To do this, we used the Python Spotipy package, which validated my profile with my playlists and recommended songs based on the data it received.

In [5]:
```
spotify_df = df.copy()
```

In [6]:
```
spotify_df.shape
```

Out[6]:
```
(156410, 27)
```

In [7]:
```
float_cols = spotify_df.dtypes[spotify_df.dtypes == 'float64'].index.values
```

In [8]:
```
ohe_cols = 'popularity'
```

In [9]:
```
# create 5 point buckets for popularity
spotify_df['bucket_popularity'] = spotify_df['popularity'].apply(lambda x: int(x/5))
```

In [10]:
```
spotify_df['consolidates_genre_lists_upd'] = spotify_df['consolidates_genre_lists'].apply(lambda x: [re.sub(' ','_',i) for i in re.findall(r"'([^']*)'"
```

In [11]:
```
spotify_df.head()
```

Out[11]:

	Unnamed: 0	Unnamed: 0.1	valence	year	acousticness	artists	danceability	duration_ms	energy	explicit	id	instrumentalness	key	liveness	loudness	mc
0	0	0	0.177	1989	0.568	['조정현']	0.447	237688	0.2150	0	2ghebdwe2pNXT4eL34T7pW	0.000001	10	0.0649	-16.478	
1	1	1	0.352	1992	0.381	['黑豹']	0.353	316160	0.6860	0	3KIuCzckjdeeVuswPo20mC	0.000000	11	0.0568	-9.103	

Fig. 4. Spotipy library used for recommendations

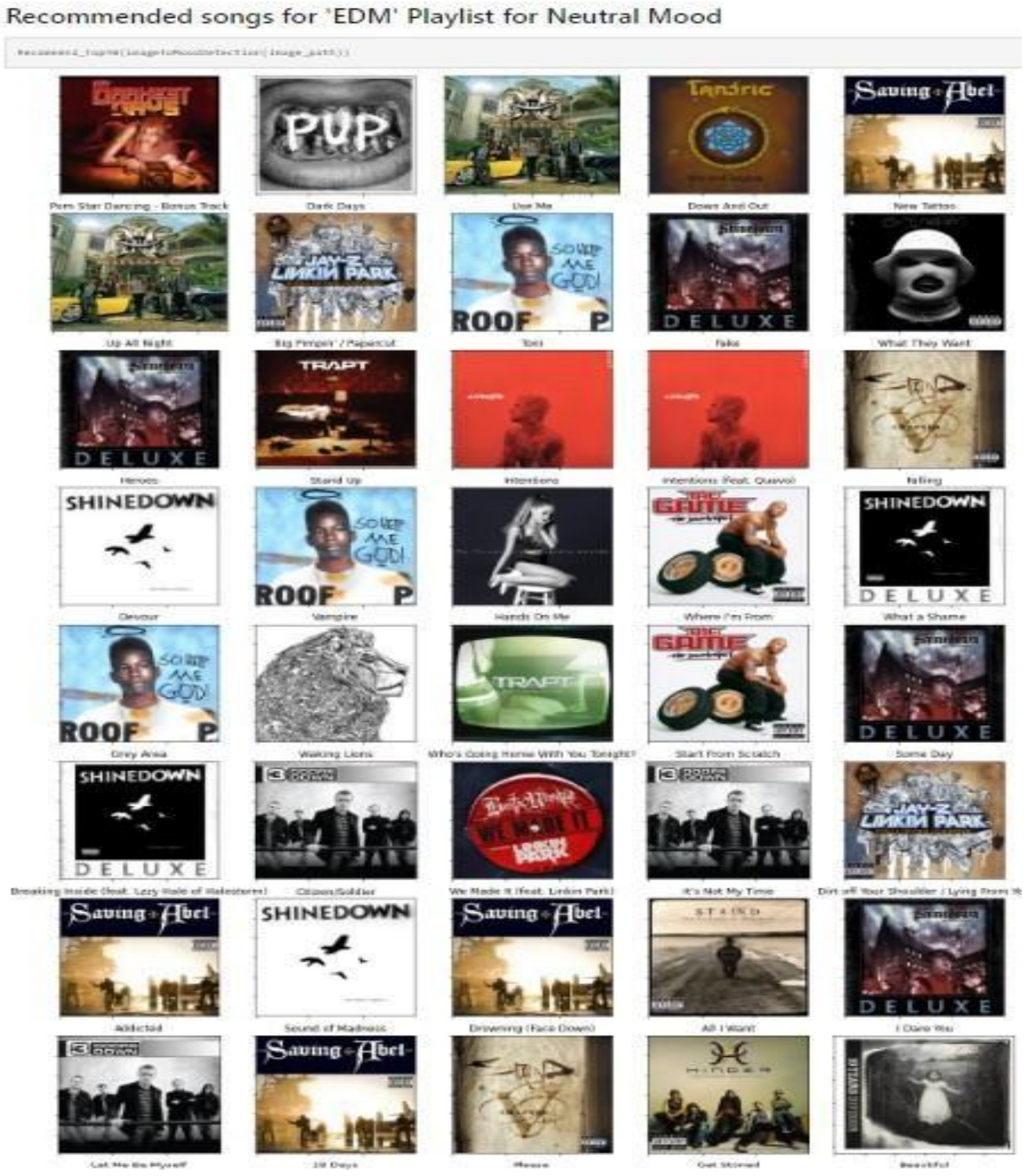


Fig. 5. Recommended songs after training

NOVELTY OF OUR PROJECT

Today we are living in a generation that we depend a lot on website-based applications for any simple tasks, e.g., booking hotels, flight, movie tickets, watching movies (Netflix etc.), shopping (Amazon, Flipkart, Myntra etc.), listening to music (Gaana etc.), advertisements, finding tourist places, finding restaurants based on location etc.

The techniques we will be using in this project are machine learning with some algorithms of deep learning.

We will also be using Natural language processing for the preprocessing of our textual data accumulated from the past usage of the data by the user.

This is all what makes our project interesting to work on and singularize us among the crowd.

OUTCOME

Music can be recommended based on the album and artist information that is available. Another approach to categorise moods is by pitch and rhythm. Regrettably, this will result in predictable advice. For example, recommending songs based on artists the user is known to like isn't very useful. The application of Neural Networks has grown in popularity as a result of evolving methods. The melodies in each class are organised using an Artificial Neural Network (ANN).

A simple system for music recommendation based on face expression identification is proposed here. It suggests music based on a person's facial expressions: happy, angry, surprised, and neutral. There is room for more upgrades and improvements. Due to the lopsided nature of each element set, progressive effective techniques to incorporate diverse highlights and functionality should be investigated in all circumstance. It is also apparent that the informational collection used to develop the grouping model might be enhanced further to improve the accuracy of the arrangement framework.

CONCLUSION

A basic framework is proposed here for the music suggestion utilizing face feeling acknowledgment. It proposes music by extracting different facial feeling of an individual: Happy, outrage, shock, impartial. There is a degree for additional upgrades and improvements. Logically compelling ways to deal with integrate various features and functionalities ought to, regardless, be examined because of the unbalanced idea of every component set. It is also seen that to improve the exactness of the game plan structure the enlightening collection used to construct the gathering model could be extended further.

The proposed framework is relying upon the identified inclination will introduce the correspondent music playlist. Since we have four feelings, we also have four playlists that offer music cuts that are painstakingly picked. For cheerful inclination, the classical music playlist will be enacted, while the new age music playlist is dedicated to the normal inclination. For the pessimistic feelings, surprised and sad the architect music playlist will be empowered to improve the client mood to a better mood.

REFERENCES

1. https://www.3ciencias.com/wp-content/uploads/2020/03/art-16_special-issue_3c_tecno_Kalasalingam-Academy-1.pdf

2. https://sci-hub.se/

3. https://ieeexplore.ieee.org/document/8299738

4. https://iopscience.iop.org/article/10.1088/1757-899X/912/6/062007/pdf

5. S. Lawrence, C. L. Giles, Ah Chung Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," in IEEE Transactions on Neural Networks, vol. 8, no. 1, pp. 98–113, Jan. 1997.

6. A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch, and M. R. Wróbel, "Human-Computer systems interaction: back-grounds and applications," 3, ch. Emotion Recognition and Its Applications, Cham: Springer International Publishing, 2014, pp. 51–62.

7. R. Ferguson, "Facial recognition software comes to loyalty", RetailWire, 2018. [Online]. Available: https://www.retailwire.com/discussion/facial-recognition-software-comes-to-loyalty/?utm_source=Listrak&utm_medium=Email&utm_term=http%3a%2f%2fwww.retailwire.com%2fdiscussion%2ffacial-recognition-software-comes-to-loyalty%2f&utm_campaign=Today+on+RetailWire%3a+Alexa-enabled+glasses%3b+Kohl%27s+omni-progress%3b+Facial+recog+loyalty+-+1%2f9%2f18. [Accessed: 29- Sep- 2018].

8. G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 34-42.

9. M. Mandila and V. Gerogiannis, "The effects of music on customer behaviorand satisfaction in the region of Larissa–the cases of two coffee bars," inInternational conference on contemporary marketing issues (ICCMI), 2012.

10. S. Alizadeh and A. Fazel, "Convolutional Neural Networks for Facial Expression Recognition," arXiv preprint arXiv:1704.06756, 2017.

11. Some Other Research Papers Referred To: https://drive.google.com/drive/folders/12qgEq6aFonSTNUgpW_GjhMVcX6VsaEpr?usp=sharing