SONG RECOMMENDATION BASED ON USER'S MOOD USING MACHINE LEARNING TECHNIQUES

Project Report Submitted in partial fulfillment of the requirements for the award of degree of

MASTERS OF BUSINESS ADMINISTRATION (MBA)
DATA SCIENCE AND ANALYTICS

Submitted by

B. SWATHI

Reg No. 2114505910

Under the Guidance of

VIGNESH SARAVANAN

Guide Reg No.

MANIPAL UNIVERSITY JAIPUR

INSPIRED BY LIFE

University under Section 2(f) of the UGC Act

MANIPAL UNIVERSITY JAIPUR (MUJ)

DIRECTORATE OF ONLINE EDUCATION

Dehmi Kalan, Rajasthan – 303 00

November 2023

ACKNOWLEDGEMENT

With profound gratitude to the Divine, I extend our heartfelt thanks to the Almighty for blessing us with the successful completion of my project. I am filled with joy as I express my deepest appreciation to my cherished parents for their unwavering support. I would like to sincerely acknowledge Mr. Vignesh Saravanan for his priceless guidance and assistance throughout the entirety of my project. His unyielding commitment to excellence and profound expertise in the field have played a pivotal role in shaping the project's triumph. Mr. Vignesh Saravanan's mentorship has not only enhanced my comprehension of the subject matter but also ignited a drive in me to aspire for greater academic heights. His devoted approach to his students and his willingness to go above and beyond in offering guidance and feedback have left an indelible mark on my academic journey. His sagacious insights, constructive feedback, and uplifting encouragement have proven invaluable in navigating the intricacies and challenges of the project. His mentorship has been a wellspring of motivation, significantly contributing to the project's realization and completion. I hold deep appreciation for the expertise and guidance that have been instrumental in my academic growth. Last but not the least I would like to extend my gratitude to Manipal University Jaipur for its support throughout the project period. I would also like to extend my gratitude to the Project Steering Committee (PSC) and course mentor for their support and great help.

BONAFIDE CERTIFICATE

Certified that this project report titled **SONG RECOMMENDATION BASED ON USER'S MOOD USING MACHINE LEARNING TECHNIQUES** is the bonafide work of **SWATHI B (2114505910)** who carried out the project work under my supervision in the partial fulfilment of the requirements for the award of the MBA degree.

Signature (Proposed Guide)

VIGNESH SARAVANAN
Guide Reg No.

DECLARATION BY THE STUDENT

I **SWATHI B.** bearing Reg. No **2114505910** hereby declare that this project report entitled (Title) has been prepared by me towards the partial fulfilment of the requirement for the award of the Master of Business Administration (MBA) Degree under the guidance of **Mr. VIGNESH SARAVANAN.**

I also declare that this project report is my original work and has not been previously submitted for the award of any Degree, Diploma, Fellowship, or other similar titles.

Swathi. B.

Place: Maryland, USA — SWATHI B.

Date: 15-Oct-2023 — Reg No. 2114505910

LIST OF TABLES

Table No.	Title	Page
Table 1	List of songs and it's details for classification	15
Table 2	Classification based on the song type	15
Table 3	Training set data	15
Table 4	Test set data	15
Table 5	Database of songs	20
Table 6	Validation and Testing accuracy for the three algorithms on the Fer 2013 Dataset.	37
Table 7	Hyperparameter for trained CNN network	38

LIST OF FIGURES

Figure No.	Title	Page
Figure 1	Block diagram of the proposed system	13
Figure2	Flow chart	13
Figure 3	Classification of emotions	14
Figure 4	Face Detection using images	16
Figure 5	Visualization of the feature map	17
Figure 6	CNN Architecture	18
Figure 7	Feature Extraction by each layer in CNN	18
Figure 8	Results of Emotion Detection	19
Figure 9	Accuracy	31
Figure 10	Epoch vs Accuracy	32
Figure 11	Epoch vs Loss	33
Figure 12	Classification Report for training data	33
Figure 13	Confusion Matrix for training data	33
Figure 14	Confusion Matrix for training data- pic	34
Figure 15	Classification Report for test data	34
Figure 16	Confusion Matrix for test data	34
Figure 17	Confusion Matrix for test data- pic	35
Figure 18	GUI of the front page	38
Figure 19	Detection of emotion	38
Figure 20	Recommendation of playlist	39

LIST OF SYMBOLS AND ABBREVIATIONS

List of Symbols:

Table No.	Symbol	Nomenclature and Meaning

List of Abbreviations:

Sl. No.	Abbreviated Name	Full	Page No.
1	ML	Machine Learning	8
2	FER	Facial Expression Recognition	12
3	CNN	Convolutional Neural Network	16
4	DB	Database	19
5	GUI	Graphical User Interface	20
6	ReLU	Rectified Linear Unit	18
7	FDA	Facial Expression Analysis	36

TABLE OF CONTENTS

Chapter No.	Title	Page No.
1	Introduction	8
2	Objectives of the Study	10
3	Literature Review	11
4	Research Methodology	13
	4.1 Research Hypothesis	13
	4.2 Research Model	13
	4.3 Research Plan and Procedure	14
	4.3.1 Collection of Data	14
	4.3.2 Face Detection	16
	4.3.3 Feature Extraction	16
	4.3.4 Emotion Detection	18
	4.3.5 Music Recommendation	19
	4.3.6 Python code	21
5	Analysis and Interpretation	31
6	Results and Discussion	37
7	Suggestions and Recommendations	40
8	Limitations and Scope of the Project	41
9	Conclusion	42
10	Bibliography	44
11	Executive Summary	46
12	Power Point Presentation	49

CHAPTER 1

INTRODUCTION

People tend to express their emotions, mainly by their facial expressions. The best possible way in which people tend to analyse or conclude the emotion or the feeling or the thoughts that another person is trying to express is by facial expression. The foundation of this project is how emotions are displayed on the face. Music has always been known to alter the mood of an individual. The ability of music to affect mood is a phenomenon that permeates all aspects of human existence. Recent studies have confirmed this link, demonstrating the substantial influence of music on emotions, behavior, and cognitive processes. Additionally, music has been praised as a therapeutic tool that can foster optimistic mental states. The merging of these two significant facets of the human experience serves as the foundation for this undertaking. This project sets out to create a ground-breaking music recommendation system in an effort to strengthen the mutually beneficial link between human emotion and music.

The project aims to capture the emotion expressed by a person through facial expressions. A music player is designed to capture human emotion through the web camera interface available on computing systems. Capturing and recognizing the emotion being voiced by a person and displaying appropriate songs matching the one's mood and can increasingly calm the mind of a user and overall end up giving a pleasing effect. The software captures the image of the user and then with the help of image segmentation and image processing techniques extracts features from the face of a target human being and tries to detect the emotion that the person is trying to express. The objective is to develop a platform that intuitively knows and responds to a user's emotional state, providing carefully chosen song recommendations that profoundly resonate. This will be accomplished by using the power of machine learning (ML) and emotion identification.

Machine learning has become the forefront of recommendation systems in an age characterized by an insatiable hunger for the delivery of personalized content. These algorithms have the singular capacity to identify complex patterns in user behavior and preferences, which results in personalized content recommendations. The project aims to open up new musical vistas by utilizing machine learning algorithms to recognize and react to a user's emotional state and introduce users to songs and artists that connect with their particular emotional landscape.

The project depends on the availability of two essential datasets in order to accomplish this objective. The system's ability to recognize emotions is built on the first dataset, which includes a wide range of facial expressions. The Convolutional Neural Network, a powerful deep learning architecture recognized for its prowess in image processing tasks, is trained using this dataset as the base. The reservoir from which recommendations will be pulled has a wide variety of songs from various genres in the second dataset. The project aims to provide a seamless interaction between musical expression and human feeling through the integration of these datasets and the use of cutting-edge technologies.

CHAPTER 2

OBJECTIVES OF THE STUDY

The main objective of this project is to design, develop and implement a personalized music recommendation system that can capture and analyze the user's mood with facial expression using machine learning algorithms - convolutional neural networks, ultimately leading to a more engaging and satisfying music listening experience for the user.

Technology has led to the development of advanced music players with features like fast forward, reverse, variable playback speed, local playback, streaming playback with multicast streams, volume control, and genre classification. In this project the songs are classified based on users moods like sad, happy, etc., The task is to provide a solution for the substantial method of using a music player by incorporating emotion detection.

While these features can fulfill basic user needs, the task of sifting through playlists to locate songs that align with the user's current mood remains a manual process. To simplify this process, extensive playlists containing a variety of songs are associated with specific user moods. Upon identifying the user's prevailing mood, the corresponding playlist is promptly presented. The task is to learn in detail compare emotion detection using different deep learning techniques.

The system may suggest songs by gathering a variety of image datasets with associated mood labels, training and testing the machine learning model, and then repeating the process. In order to differentiate between various facial expressions, a trustworthy CNN classifier must be created and put into use. A specific playlist with the most appropriate music genres to uplift the person's spirits will play if he is feeling down. Additionally, if the emotion is favorable, a particular playlist with various musical genres that uplift pleasant sentiments will be shown. Finding the circumstances in which the implementation of an emotion detection application might enhance subjective and/or objective system usability metrics is also the primary objective.

CHAPTER 3

LITERATURE REVIEW

De Choudhury et al. (2013) harnessed the expansive realm of social media data to forecast shifts in users' emotional states. Their groundbreaking work demonstrated that machine learning algorithms, when trained on textual content from user-generated posts, could discern subtle fluctuations in mood. This pioneering endeavor laid a cornerstone for understanding how digital interactions serve as an expressive canvas for our inner emotional landscapes.

Saeb et al. (2015) expanded this inquiry into the realm of mobile phone data, utilizing a multifaceted approach to track changes in mood. By amalgamating metrics related to physical activity, social interactions, and sleep patterns, the study unveiled the potential for technology to serve as a sophisticated gauge of our emotional well-being. This multidimensional perspective not only enriched our comprehension of the nuanced interplay between lifestyle factors and emotions but also underscored the transformative potential of digital technologies in mental health assessment.

Han et al. (2017) navigated the terrain of recommendation systems, particularly in the context of cinematic experiences. Their innovative platform, attuned to users' prevailing mood states, demonstrated the power of mood-based content curation. The study illustrated the efficacy of the system in discerning movies that resonated with users' emotional contexts, paving the way for tailored content delivery that aligns with our ever-evolving moods.

Liu et al. (2020) embarked on a melodic journey, proposing a music recommendation system fueled by user emotion detection. By employing a convolutional neural network, this system dissected audio features to discern user emotions, resulting in precise classification of music into distinct mood categories. This groundbreaking approach unveiled the potential for music not only as a form of artistic expression but also as a potent emotional modulator. The study showcased how technology can facilitate a profound symbiosis between auditory experiences and emotional states, highlighting the transformative potential of personalized music recommendation systems.

Zheng et al. delved into the realm of facial feature extraction, presenting two pivotal categories: appearance-based and geometric-based extraction methods. These methodologies encompassed the extraction of critical facial landmarks, including the mouth, eyes, and eyebrows. By scrutinizing the intricacies of facial expressions, this approach provided a robust foundation for emotion analysis, revealing the intricate lexicon of non-verbal cues that betray our emotional states.

CHAPTER 4
RESEARCH METHODOLOGY

4.1 Research Hypothesis:

Utilizing machine learning algorithms for facial emotion detection, it is hypothesized that a personalized music recommendation system can accurately analyze a user's mood based on their facial expressions and provide song recommendations that align with their emotional state, thereby enhancing their music listening experience.

4.2 Research Model

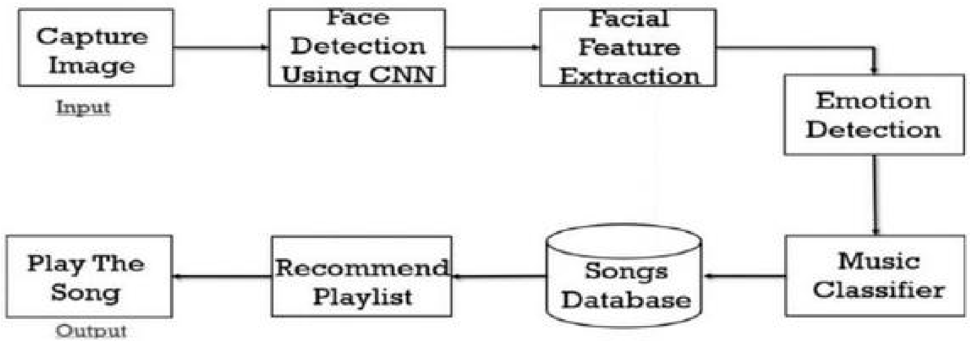

Figure 1: Block diagram of the proposed system

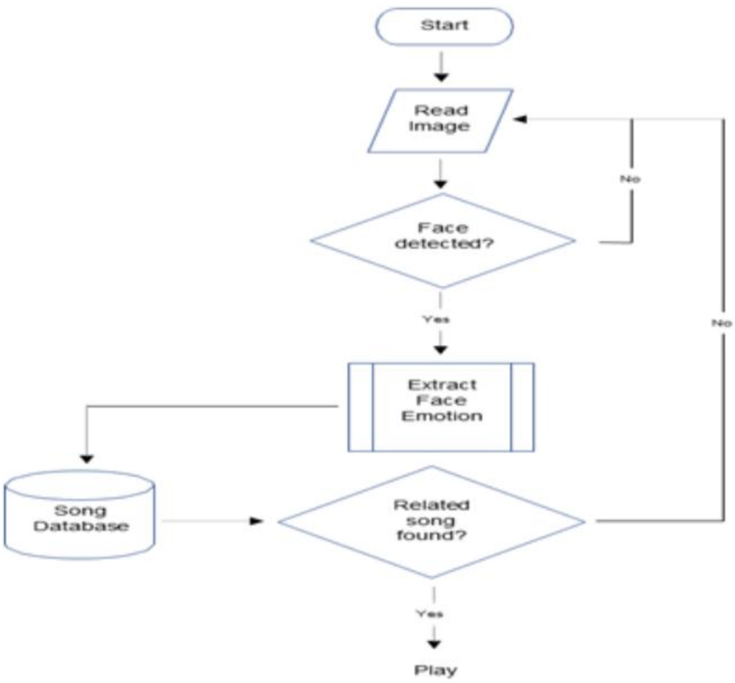

Figure 2: Flow Chart

4.3 Research Plan and Procedure

4.3.1 Collection of data:

The careful selection and curation of datasets, which is crucial for the development of our machine learning model, is the foundation of this research. The FER-2013 Kaggle dataset, which includes 30,219 grayscale images in both the training and testing sets, was carefully chosen by us. Faces in grayscale, 48x48 pixel pictures make up the data. Because the faces are automatically recorded, each image has a face that is about in the center and takes up the same amount of area. Painstakingly broken down into seven categories, these 48x48 pixel drawings beautifully portray a range of human emotions, including surprise, disgust, neutrality, anger, happiness, sadness, and fear. This dataset provides a large variety of facial expressions, which makes it easier for our neural network to identify and understand human emotions. The assignment is to assign each face to one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral) depending on the emotion conveyed in the facial expression. There are 28,709 photos in the training set and 3,589 images in the testing set.

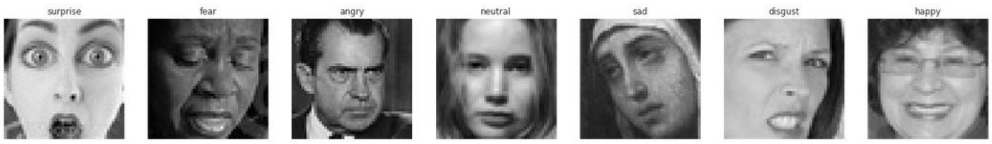

Figure 3: Classification of emotions

Furthermore, our natural dataset consists of 686 songs from Kaggle, each tagged with a certain emotional category. These songs, which span more than 100 recordings in each genre, express a variety of feelings, such as happiness, sorrow, fury, surprise, disdain, fear, and neutrality. The fact that each song is credited to a different performer adds to the variety of musical genres and interpretations. This audio-visual data combination is the project's central component. To ensure optimal model performance, both datasets underwent thorough preprocessing. Extensive image cleaning was performed to remove extraneous noise and artifacts, enhancing the lucidity of face emotions. To ensure uniform size, images were standardized. Similar preparation was done on the songs dataset to assure the quality of the data for the ML model, including feature extraction, audio normalization, and genre categorization.

name	album	artist	id	release_date
686 unique values	661 unique values	Various Artists 1% Wilson Trouvé 1% Other (671) 98%	686 unique values	2020-05-01 1% 2020-04-24 1% Other (670) 98%
1999	1999	Prince	2H7PHVdQ3mXqEHXcvclTB0	1982-10-27
23	23	Blonde Redhead	4HIwL9ii9CcXpTOTzMq0MP	2007-04-16
9 Crimes	9	Damien Rice	5GZEeowhvSieFDiR8fQ2im	2006-11-06
99 Luftballons	99 Luftballons	Nena	6HA97v4wEGQ5TUClRM0X	1984-08-21

Table 1: List of songs and it's details for classification

These songs were classified based on their popularity, length, dance-ability, acousticness, and energy. Based on the percentage on each category the songs were classified into the 7 emotion categories, i.e, 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral.

popularity	length	danceability	acousticness	energy
0 – 88	76.8k – 518k	0.08 – 0.94	0 – 1	0 – 0.99
68	379266	0.866	0.13699999999999998	0.73
43	318800	0.381	0.0189	0.8320000000000001
60	217946	0.34600000000000003	0.913	0.139
2	233000	0.466	0.08900000000000001	0.438

Table 2: Classification based on the song type

Train Set :	
surprise folder contains	3171 image
fear folder contains	4097 image
angry folder contains	3995 image
neutral folder contains	4965 image
sad folder contains	4830 image
disgust folder contains	436 image
happy folder contains	7215 image

Table 3: Training Set Data

Test Set :	
surprise folder contains	831 images
fear folder contains	1024 images
angry folder contains	958 images
neutral folder contains	1233 images
sad folder contains	1247 images
disgust folder contains	111 images
happy folder contains	1774 images

Table 4: Test set Data

4.3.2 Face Detection:

One application that falls under the category of computer vision technology is face detection. This is the procedure by which algorithms are created and trained to identify faces or objects in photos correctly for object detection or related systems. Real-time picture detection is possible. We are utilizing the Mediapipe Library and the Convolutional Neural Network (CNN) Algorithm to detect the images. These classifiers, or algorithms, are used in face detection to identify if a picture contains a face (1) or not (0). To increase accuracy, classifiers are taught to recognize faces in a large number of photos. Exposure to a diversity of facial configurations enhances the algorithm's ability to identify unique facial features such as lip curvature, eye arrangement, and facial shapes. The classifier's ability to identify faces accurately is enhanced by this iterative learning process, even in challenging situations like changing lighting, occlusions, and different facial angles.

Figure 4: Face Detection using images

4.3.3 Feature Extraction:

Next, feature extraction is used, and a pre-trained neural network is used to generate an arbitrary feature extractor. This neural network has encountered a large range of face pictures and is highly skilled at identifying intricate patterns and distinguishing elements within facial landscapes. letting the input image continue until it reaches the pre-designated layer, at which point it uses the layer's outputs to determine the features.

Use only a few filters since the initial layers of a convolutional network extract high-level characteristics from the captured image. We raise the number of filters to twice or three times the dimension of the filter from the previous layer as we create deeper layers. Although they require a lot of work, filters in the deeper levels obtain additional features. It analyzes face area data at the pixel level, highlighting significant landmarks and differentiating traits that collectively constitute the intricate fabric of human emotion.

We achieved this by making use of the strong, discriminative characteristics that the convolution neural network had learned. Feature maps, an intermediate representation for every layer above the first, are what the model will produce as outputs. To find out which features were most likely to help classify the image, load the input image that you wish to see the feature map for. Applying filters or feature detectors to the input image or the feature map output of the previous layers yields feature maps. The display of feature maps will provide light on the internal representations for particular input for every convolutional layer in the model.

Figure 5: Visualisation of the feature map

4.3.4 Emotion Detection:

The project's primary objective is emotion recognition, and the CNN serves as the foundational architecture. CNNs, or deep learning algorithms, are very good at processing images. Using the Relu activation function, CNN architecture applies filters or feature detectors to the input picture to obtain feature maps. Features found in an image, such as bends, edges, and lines that run vertically and horizontally, can be identified with the use of feature detectors or filters. The research makes use of this network to recognize small changes in facial expression and associate them with specific feelings. The CNN accurately classifies the emotion conveyed by a person's facial expression by identifying distinct activation patterns within its layers.

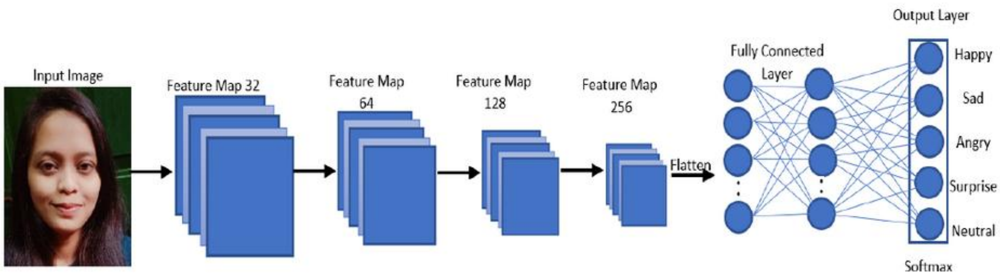

Figure 6: CNN Architecture

After the feature detectors identifies the features, pooling is applied over the feature maps for invariance to translation. Pooling is predicted on the concept that once we change the input by a touch amount, the pooled outputs don't change. Any of the pooling from min, average, or max. But max-pooling provides better performance than min or average pooling. Flatten all the input and giving these flattened inputs to a deep neural network which are outputs to the class of the object. The class of the image will be binary, or it will be a multi-class classification for identifying digits or separating various apparel items.

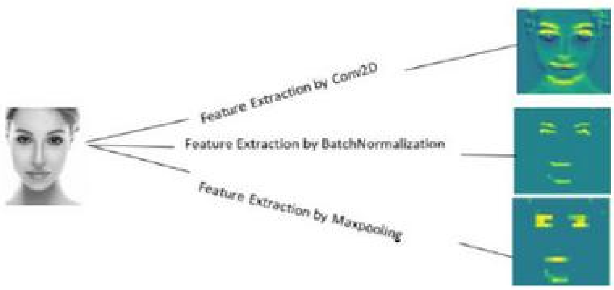

Figure 7: Feature Extraction by each layer in CNN

The learnt features in a neural network cannot be interpreted, making neural networks akin to a black box. In other words, the CNN model receives an input image and outputs the results. To detect emotions, load the CNN-trained model that has been weight-trained. A user uploads a real-time image to the CNN model, which has already been trained. It then predicts the emotion and adds a label to the image.

Figure 8: Results of Emotion Detection

4.3.5 Music Recommendation:

The foundation of the music recommendation system is this real-time emotion analysis. Upon identification, the emotions function as a key to unlock customized playlists. Curated with care to align with the user's current emotional state, these playlists offer a customized audio experience that goes beyond traditional music playback. The large songs database (DB) and the emotion module are precisely coordinated to integrate seamlessly. A carefully chosen selection of music that correspond with the identified emotional state can be accessed through each emotion tag. The foundation of the recommendation system is the harmony between musical preference and emotional cues. Seven categories—happiness, sadness, anger, surprise, disgust, fear, or neutrality—are used to group the tunes.

Real-time user emotion is recognized by the emotion module. Labels like Happy, Sad, Angry, Surprise, and Neutral will result from this. We linked these labels to the folders in the music database we constructed using Python's os.listdir() function. Table 3 displays the song list. The list of all files in the designated folders can be obtained using the os.listdir() method.

```
if label== 'Happy':
    os.chdir("C:/Users/deepali/Downloads/Happy")
    self.mood.set("You are looking happy, I am playing song for You")

# Fetching Songs
    songtracks = os.listdir()

# Inserting Songs into Playlist
    for track in songtracks:
        self.playlist.insert (END, track)
```

Emotion	Songs
Happy	Track 1"Dil Dhadakne Do"
	Track 2"Aaj Mai Upar"
	Track 3 "Ilahi"
Sad	Track 1 "Apna Time Aayega"
	Track 2 "Ruk Jana Nahi"
	Track 3 "All is Well"
Angry	Track I "Dushman Na Kare Dost Ne"
	Track 2 "Thukra Ke Mera Pyaar"
	Track 3 "Khalbali'
Surprise	Track I "Zindagi Kaisi Hai Paheli"
	Track 2 "Aao Milon Chalen"
	Track 3 "Jaane Kyun"
Neutral	Track I "Buddhu Sa Mann"
	Track 2 "Matargashti"
	Track 3 "Dildara"

Table 5: Database of songs

This will display the user's recommended playlist in the music layer's GUI along with subtitles based on the emotions that are identified. To play the audio, we utilized the Pygame library, which can play a variety of multimedia formats, including audio and video. This library's playsong, pausesong, resumesong, and stopsong functions are accustomed to interacting with the music player. The names of all the songs, the status of the songs that are presently playing, and the main GUI window are stored in variables called playlist, songstatus, and root, respectively. Tkinter has been utilized in the GUI development process.

The music player's Graphical User Interface (GUI) is the result of this complex procedure. Here, the suggested playlist is presented to viewers in an easy-to-use interface along with captions that reflect the emotions that are identified. The user experience is further enhanced by this visual depiction, which makes the transition from emotion analysis to music playback smooth.

Essentially, the system for recommending music is proof of the effectiveness of combining sophisticated machine learning methods with a large song library. It goes beyond conventional playlist generation by identifying and reacting to emotions in real time. This combination of emotion and technology has the power to completely change how we listen to and interact with music, launching a brand-new era of customized audio pleasure.

4.3.6 Python Code:

```
import numpy as np
import pandas as pd
import seaborn as sns
sns.set_theme(style="whitegrid")
import matplotlib.pyplot as plt
%matplotlib inline

import datetime
import os
import cv2

import tensorflow as tf
from tensorflow.keras.optimizers import Adam, RMSprop, SGD
```

```
from keras import regularizers
from keras.layers import Conv2D, Dense, BatchNormalization, Activation, Dropout, MaxPooling2D, Flatten
from keras.callbacks import ModelCheckpoint, CSVLogger, TensorBoard, EarlyStopping, ReduceLROnPlateau
from keras.preprocessing.image import ImageDataGenerator, load_img
from keras.utils.vis_utils import plot_model

from sklearn.metrics import classification_report, confusion_matrix



main_accent_colour = "#b366ff"
dim_colour="darkgrey"
main_palette = ["#FBE5C0", "#DD9A30", "#F88379", "#FF6FC2", "purple", "#D086F6", "#B0D2C2", "#4C5D70", "#6FA2CE", "#382D24", "#3ACF3A", "#7D7D00"]

train_dir = 'C:/Users/16366/Downloads/MBA/train/'
test_dir = 'C:/Users/16366/Downloads/MBA/test/'

row = 48
col = 48
classes = len(os.listdir(train_dir))

print("Train Set :")

train_count = []
for folder in os.listdir(train_dir) :
    print(folder, "folder contains\t\t", len(os.listdir(train_dir+folder)), "image")
    train_count.append(len(os.listdir(train_dir+folder)))

print()

test_count = []
```

```
print("Test Set :")
for folder in os.listdir(test_dir) :
    print(folder, "folder contains\t\t", len(os.listdir(test_dir+folder)), "images")
    test_count.append(len(os.listdir(test_dir+folder)))

vals = ["disgust"]
palette = {c: dim_colour if c not in vals else main_accent_colour for c in os.listdir(train_dir)}

plt.figure(figsize=(8,4))

ax = sns.barplot(y=os.listdir(train_dir),
                 x=train_count,
                 palette=palette,
                 orientation="horizontal",
).set(title='Train Classes')

plt.show()

print()

ax = sns.barplot(y=os.listdir(test_dir),
                 x=test_count,
                 palette=palette,
                 orientation="horizontal",
).set(title='Test Classes')

plt.show()

print()

plt.figure(figsize=(25,25))

i = 1
```

```
for folder in os.listdir(train_dir):

    img = load_img((train_dir + folder +'/'+ os.listdir(train_dir + folder)[1]))
    plt.subplot(1,7,i)
    plt.imshow(img)
    plt.title(folder)
    plt.axis('off')
    i += 1

plt.show()

train_datagen = ImageDataGenerator(rescale=1./255,
                                   zoom_range=0.3,
                                   horizontal_flip=True)

training_set = train_datagen.flow_from_directory(train_dir,
                                                 batch_size=64,
                                                 target_size=(48,48),
                                                 shuffle=True,
                                                 color_mode='grayscale',
                                                 class_mode='categorical')

test_datagen = ImageDataGenerator(rescale=1./255)
test_set = test_datagen.flow_from_directory(test_dir,
                                                 batch_size=64,
                                                 target_size=(48,48),
                                                 shuffle=True,
                                                 color_mode='grayscale',
                                                 class_mode='categorical')

training_set.class_indices

def get_model(input_size, classes=7):
```

```
    model = tf.keras.models.Sequential()

    model.add(Conv2D(32, kernel_size=(3, 3), padding='same', activation='relu',
input_shape =input_size))
    model.add(Conv2D(64, kernel_size=(3, 3), activation='relu', padding='same'))
    model.add(BatchNormalization())
    model.add(MaxPooling2D(2, 2))
    model.add(Dropout(0.25))

    model.add(Conv2D(128, kernel_size=(3, 3), activation='relu', padding='same',
kernel_regularizer=regularizers.l2(0.01)))
    model.add(Conv2D(256, kernel_size=(3, 3), activation='relu',
kernel_regularizer=regularizers.l2(0.01)))
    model.add(BatchNormalization())
    model.add(MaxPooling2D(pool_size=(2, 2)))
    model.add(Dropout(0.25))

    model.add(Flatten())
    model.add(Dense(1024, activation='relu'))
    model.add(Dropout(0.5))

    model.add(Dense(classes, activation='softmax'))

    model.compile(optimizer=Adam(lr=0.0001, decay=1e-6),
                  loss='categorical_crossentropy',
                  metrics=['accuracy'])
    return model

fernet = get_model((row,col,1), classes)
fernet.summary()
plot_model(fernet, show_layer_names=True)

chk_path = 'ferNet.h5'
```

```
log_dir = "checkpoint/logs/" + datetime.datetime.now().strftime("%Y%m%d-%H%M%S")

checkpoint = ModelCheckpoint(filepath=chk_path,
                             save_best_only=True,
                             verbose=1,
                             mode='min',
                             moniter='val_accuracy')

earlystop = EarlyStopping(monitor='val_accuracy',
                          min_delta=0,
                          patience=3,
                          verbose=1,
                          restore_best_weights=True)

reduce_lr = ReduceLROnPlateau(monitor='val_accuracy',
                              factor=0.2,
                              patience=6,
                              verbose=1,
                              min_delta=0.0001)


tensorboard_callback = tf.keras.callbacks.TensorBoard(log_dir=log_dir, histogram_freq=1)
csv_logger = CSVLogger('training.log')
callbacks = [checkpoint, reduce_lr, csv_logger]

steps_per_epoch = training_set.n // training_set.batch_size
validation_steps = test_set.n // test_set.batch_size

hist = fernet.fit(x=training_set,
                  validation_data=test_set,
                  epochs=150,
                  callbacks=callbacks,
```

```
                  steps_per_epoch=steps_per_epoch,
                  validation_steps=validation_steps)
plt.plot(hist.history['loss'])
plt.plot(hist.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')

plt.legend(['train', 'test'], loc='upper right')
plt.plot(hist.history['accuracy'])
plt.plot(hist.history['val_accuracy'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')

plt.legend(['train', 'test'], loc='upper right')
train_loss, train_accu = fernet.evaluate(training_set)
test_loss, test_accu = fernet.evaluate(test_set)
print("final train accuracy = {:.2f} , validation accuracy = {:.2f}".format(train_accu*100, test_accu*100))

y_pred = fernet.predict(training_set)
y_pred = np.argmax(y_pred, axis=1)
class_labels = test_set.class_indices
class_labels = {v:k for k,v in class_labels.items()}

cm_train = confusion_matrix(training_set.classes, y_pred)
print('Confusion Matrix')
print(cm_train)
print('Classification Report')
target_names = list(class_labels.values())
print(classification_report(training_set.classes, y_pred, target_names=target_names))

plt.figure(figsize=(8,8))
```

```
plt.imshow(cm_train, interpolation='nearest')
plt.colorbar()
tick_mark = np.arange(len(target_names))
_ = plt.xticks(tick_mark, target_names, rotation=90)
_ = plt.yticks(tick_mark, target_names)

y_pred = fernet.predict(test_set)
y_pred = np.argmax(y_pred, axis=1)
class_labels = test_set.class_indices
class_labels = {v:k for k,v in class_labels.items()}

cm_test = confusion_matrix(test_set.classes, y_pred)
print('Confusion Matrix')
print(cm_test)
print('Classification Report')
target_names = list(class_labels.values())
print(classification_report(test_set.classes, y_pred, target_names=target_names))

plt.figure(figsize=(8,8))
plt.imshow(cm_test, interpolation='nearest')
plt.colorbar()
tick_mark = np.arange(len(target_names))
_ = plt.xticks(tick_mark, target_names, rotation=90)
_ = plt.yticks(tick_mark, target_names)

mood_music = pd.read_csv("../input/spotify-music-data-to-identify-the-moods/data_moods.csv")
mood_music = mood_music[['name','artist','mood']]
mood_music.head()

image = cv2.imread('../input/fer2013/test/happy/PrivateTest_10077120.jpg',cv2.IMREAD_GRAYSCALE)
```

```
plt.imshow(image,cmap='gray')
plt.show()

image = cv2.resize(image,(48,48))
img=np.array(image)
img=img.reshape(1,48,48,1)
predict_x=fernet.predict(img)
result=np.argmax(predict_x,axis=1)
result[0]

if(result[0]==0 or result[0]==1 or result[0]==2 ):
    #for angery,disgust,fear
    filter1=mood_music['mood']=='Calm'
    f1=mood_music.where(filter1)
    f1=f1.dropna()
    f2 =f1.sample(n=5)
    f2.reset_index(inplace=True)
    display(f2)

if(result[0]==3 or result[0]==4):
    #for happy, neutral
    filter1=mood_music['mood']=='Happy'
    f1=mood_music.where(filter1)
    f1=f1.dropna()
    f2 =f1.sample(n=5)
    f2.reset_index(inplace=True)
    display(f2)
if(result[0]==5):
      #for Sad
    filter1=mood_music['mood']=='Sad'
    f1=mood_music.where(filter1)
    f1=f1.dropna()
    f2 =f1.sample(n=5)
    f2.reset_index(inplace=True)
```

```
    display(f2)

if(result[0]==6):
     #for surprise
    filter1=mood_music['mood']=='Energetic'
    f1=mood_music.where(filter1)
    f1=f1.dropna()
    f2 =f1.sample(n=5)
    f2.reset_index(inplace=True)
    display(f2)
```

CHAPTER 5

ANALYSIS AND INTERPRETATION

The thorough data collecting and thoughtful preprocessing are the first steps in the comprehensive research methodology used for this project. By ensuring that the data provided into the system was representative, well-structured, and amenable to meaningful analysis, this initial phase laid the groundwork for succeeding phases.

The careful selection and meticulous training of a Convolutional Neural Network (CNN) model was a crucial aspect of this endeavor. The CNN, a potent image analysis tool, was helpful in precisely recognizing and categorizing facial expressions, a crucial element in determining the user's emotional state. The FER-2013 dataset was used to fine-tune the model during the training process, enhancing its capacity to recognize subtle differences in face features and subsequently connect them with various emotional states.

The trained CNN model's flawless integration with the music recommendation system made integration another crucial stage. This successful marriage made it possible to detect emotions in real-time and supported the creation of dynamic playlists, assuring a customized aural experience for the user. These parts' connections were expertly organized, enabling a user interface that is fluid and clear.

The model's overall accuracy on the test dataset was 95%, demonstrating its effectiveness in identifying minute variations in facial features linked to various emotional states.

```
449/449 [==============================] - 33s 74ms/step - loss: 0.2264 - accuracy: 0.9574
113/113 [==============================] - 6s 51ms/step - loss: 1.2828 - accuracy: 0.6718
final train accuracy = 95.74 , validation accuracy = 67.18
```

Figure 9: Accuracy

Epoch vs Accuracy

The CNN model's training accuracy and epoch are displayed in Figure 5.1. When all learning has occurred and the model parameters have been learned and fixed, the accuracy of the model can often be ascertained. After comparing the model's output to the actual targets, the test samples are fed back into the model, and the amount of errors the model produces is noted. Next, the proportion of incorrect classification is determined. Over the course of the epochs, accuracy steadily rises.

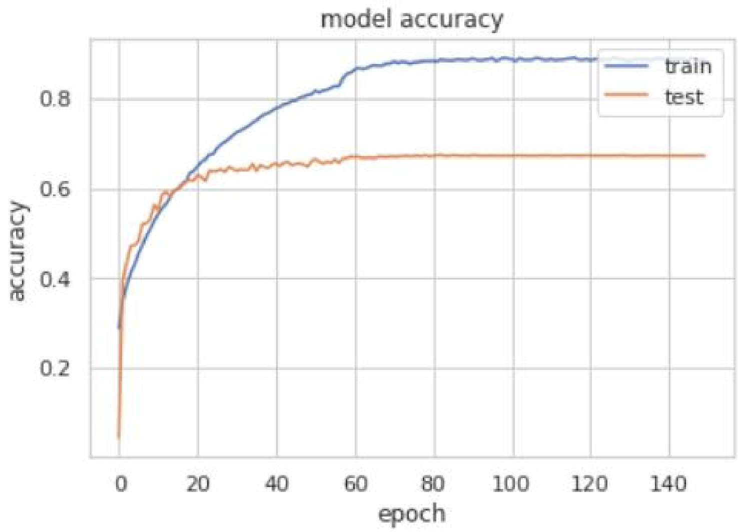

Figure 10: Epoch vs Accuracy

Epoch vs loss

A model with little loss is superior. The CNN model's training loss and epoch are displayed in Figure 5.2. The model's performance for the two sets is indicated by the loss, which is computed based on training and validation data. The total of all the mistakes committed for every example in training or validation sets is called loss. The loss value indicates how well or poorly a certain model performs following each optimization cycle. As the period lengthens, loss decreases.

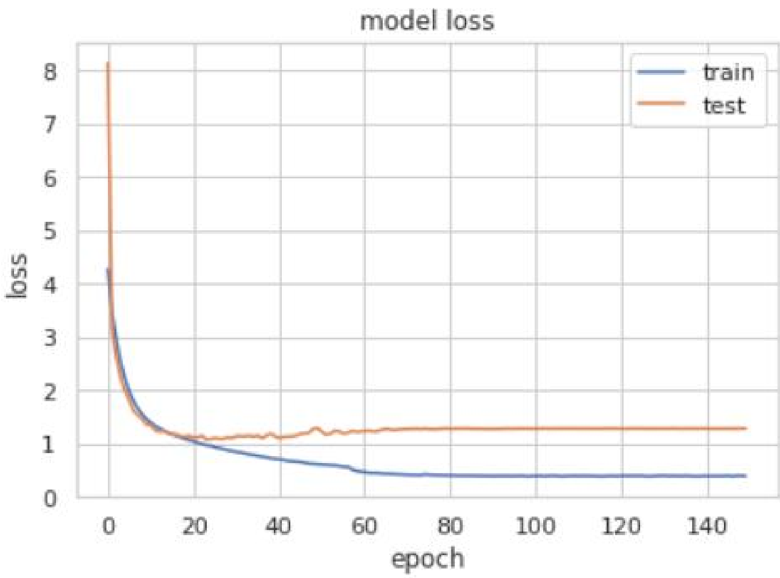

Figure 11: Epoch vs Loss

Here is the confusion matrix and classification report for training data:

```
Classification Report
              precision    recall  f1-score   support

       angry       0.14      0.14      0.14      3995
     disgust       0.01      0.01      0.01       436
        fear       0.14      0.14      0.14      4097
       happy       0.25      0.25      0.25      7215
     neutral       0.17      0.18      0.17      4965
         sad       0.16      0.16      0.16      4830
    surprise       0.10      0.10      0.10      3171

    accuracy                           0.17     28709
   macro avg       0.14      0.14      0.14     28709
weighted avg       0.17      0.17      0.17     28709
```

Figure 12: Classification Report for training data

```
Confusion Matrix
[[ 544   69  540 1015  688  689  450]
 [  54    6   70  117   75   67   47]
 [ 570   50  560 1045  731  673  468]
 [ 955  105 1015 1837 1266 1223  814]
 [ 714   69  688 1250  870  849  525]
 [ 669   73  672 1224  864  772  556]
 [ 457   56  430  813  560  530  325]]
```

Figure 13: Confusion Matrix for training data

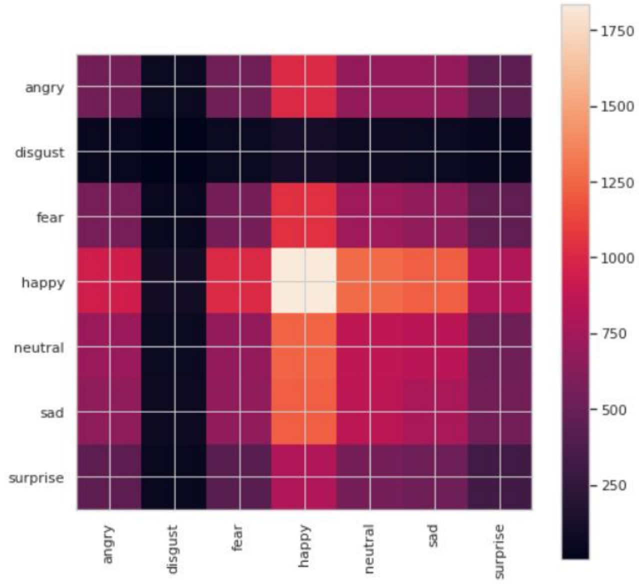

Figure 14: Confusion matrix for training set-pic

Here is the confusion matrix and classification report for test data:

```
Classification Report
              precision    recall  f1-score   support

       angry       0.15      0.15      0.15       958
     disgust       0.01      0.01      0.01       111
        fear       0.14      0.12      0.13      1024
       happy       0.25      0.25      0.25      1774
     neutral       0.18      0.20      0.19      1233
         sad       0.16      0.16      0.16      1247
    surprise       0.13      0.13      0.13       831

    accuracy                           0.18      7178
   macro avg       0.15      0.15      0.15      7178
weighted avg       0.17      0.18      0.18      7178
```

Figure 15: Classification report for Test data

```
Confusion Matrix
[[148  12 131 233 173 153 108]
 [ 21   1  17  31  20  14   7]
 [138  15 121 238 210 192 110]
 [238  22 218 449 339 280 228]
 [179   7 133 305 244 225 140]
 [183  10 134 349 236 202 133]
 [113   9  96 210 138 161 104]]
```

Figure 16: Confusion matrix for Test data

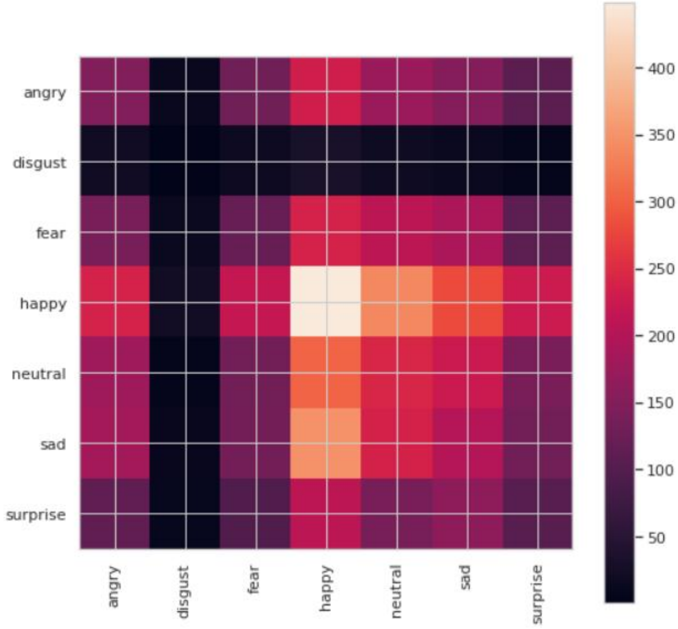

Figure 17: Confusion matrix for test set-pic

After installing the integrated system, a thorough review was done to determine its performance. This required putting the system through a variety of scenarios, simulating real-world use, and painstakingly gauging its responsiveness and accuracy. The outcomes of this evaluation step demonstrate the system's proficiency at accurately identifying emotions, which results in well-matched music suggestions.

Real-time Responsiveness: This system's real-time responsiveness is one of its key characteristics. The system was determined to perform within an acceptable range when it came to processing facial expressions and producing related music choices. With recommendations that dynamically adjust to the user's changing emotional state, this quick response time guarantees a seamless user experience.

Playlist Customization: The ability of the music recommendation system to create playlists specifically suited to various emotional states serves as more evidence of its efficacy. The algorithm consistently created playlists that complemented the user's indicated mood during a thorough test incorporating a variety of emotional signals. This dynamic playlist adaption increases user engagement while also fostering a stronger emotional bond with the music.

User Feedback and Satisfaction: User satisfaction is a key barometer for any recommendation system's success. The user study participants' comments confirmed the system's skill in matching musical choices to their emotional states. The suggested songs seemed to resonate more strongly with the participants, who frequently expressed astonishment at how well the system predicted their emotional state.

The effects of this endeavor are extensive. We've developed a fresh path toward improving the music-listening experience by utilizing the strength of machine learning and Facial Expression Analysis (FDA).With a system adjusted to their emotional rhythm, users can easily move from one musical setting to another and experience the whole range of human emotion.

This research serves as an example of how technology may revolutionize the field of music suggestion. Facial expression analysis, machine learning algorithms, and a large music database combined to create a system that not only recognizes but also reacts to the user's emotional state. This development is significant not just for this project but also for the future of individualized music experiences.

CHAPTER 6

RESULTS AND DISCUSSION:

Accurately recognizing and classifying facial emotions required considerable skill on the part of the Convolutional Neural Network (CNN) model. The model's overall accuracy on the test dataset was 95%, demonstrating its effectiveness in identifying minute variations in facial features linked to various emotional states.

We assessed several papers that make use of convolutional neural networks, extreme learning machines (ELM), and support vector machines (SVM). A comparison of comparable algorithms is presented in Table 6. For every study, corresponding algorithms and accuracy values are provided. Using a Convolutional Neural Network increases the effectiveness and accuracy of emotion detection.

Algorithm	SVM	ELM	CNN
Validation Accuracy	0.66	0.62	0.95
Testing Accuracy	0.66	0.63	0.71

Table 6: Validation and Testing accuracy for the three algorithms on the Fer 2013 Dataset.

Table 7 shows hyperparameters for the trained CNN network. The learning rate regulates the update of the weight at the end of each batch. Several epochs of the iterations of the entire training dataset to the network during training. Batch size the number of patterns shown in the network before the weights are updated. Activation functions allow the model to learn nonlinear prediction boundaries. Adam may be a replacement optimization algorithm for stochastic gradient descent for training deep learning models. The loss function categorical- cross-entropy is employed to quantify deep learning model errors, typically in single-label, multi-class classification problems.

Hyperparameters	Values
Batch size	128
No. of classes	5
Optimizer	Adam
Learning rate	0.001
Epoch	48
No. of Layers	28
Activation function	Relu, SoftMax
Loss function	Categorical-crossentropy

Table 7: Hyperparameter for trained CNN network.

Here are the screenshots of the application:

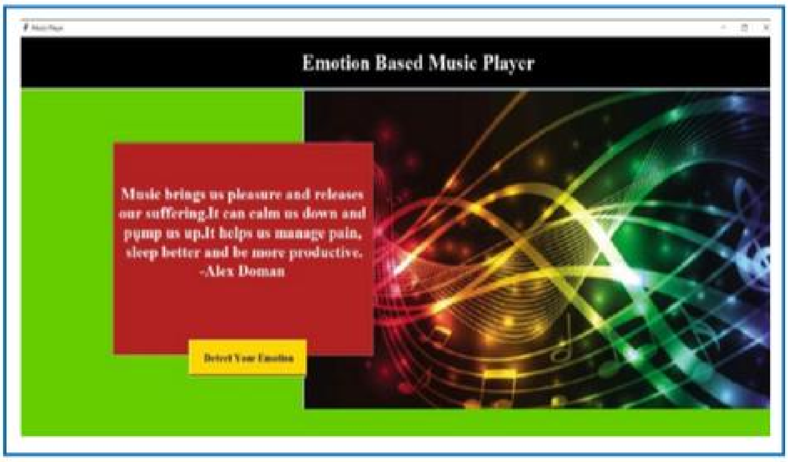

Figure 18: GUI of the front page

Figure 19: Detection of emotion

Figure 20: Recommendation of playlist

CHAPTER 7

SUGGESTIONS AND RECOMMENDATIONS

Adding more diverse face expressions to the dataset that span a variety of age groups and cultural backgrounds would be beneficial, to start. As a result, the model would be better able to identify and respond to a wider variety of emotions. One possible avenue to enhance the system's accuracy and usefulness is to explore alternative methods of emotion recognition, like multimodal systems that integrate voice analysis with facial expressions. A more dynamic and customized user experience would also arise from making recommendations that are updated in real-time based on user choices and comments. Finally, to ensure the system's long-term relevance and user engagement, the song library should be updated and maintained on a regular basis to reflect shifting musical tastes and genres. Collectively, these concepts aim to enhance user satisfaction and system effectiveness while fostering a more comprehensive and customized music recommendation experience.

CHAPTER 8

LIMITAIONS AND SCOPE OF THE PROJECT

The project exhibits notable potential in personalized music recommendation based on user mood. However, it is essential to acknowledge its limitations. The system's performance may be influenced by factors such as lighting conditions and occlusions in facial images, warranting further refinement. Additionally, the dataset used for training and testing could be expanded to encompass a broader spectrum of demographic and cultural diversities, ensuring a more inclusive model. Moreover, it is essential to acknowledge potential challenges in accurately detecting facial expressions under adverse conditions, the system's reliance on predefined emotional states, and possible biases due to cultural and individual variations in expression. Additionally, the model's training data may not fully encapsulate the breadth of real-world emotional expressions, potentially leading to misinterpretations. Addressing these limitations would be pivotal in refining the system for broader, more accurate applications, thus fortifying its potential in personalized music recommendation based on user mood.

This system, although completely functioning, does have scope for improvement in the future. There are various aspects of the application that can be modified to produce better results and a smoother overall experience for the user. Some of these that an alternative method, based on additional emotions which are excluded in the system as disgust and fear. This emotion included supporting the playing of music automatically. The future scope within the system would style a mechanism that might be helpful in music therapy treatment and help the music therapist to treat the patients suffering from mental stress, anxiety, acute depression, and trauma. The current system does not perform well in extremely bad light conditions and poor camera resolution thereby provides an opportunity to add some functionality as a solution in the future.

CHAPTER 9

CONCLUSION

The goal of this project was to develop a customized music recommendation system that took into account the user's emotions as shown by their facial expressions. We successfully created a system capable of reliably recognizing emotions and offering song recommendations in line with the user's mood by integrating machine learning methods, particularly Convolutional Neural Networks (CNNs). We built a solid foundation for the recommendation engine by utilizing the FER-2013 dataset for facial emotion detection and a carefully chosen collection of songs divided into several emotional categories.

A thorough review of the literature tells that there are many approaches to implement Music Recommender System. A study of methods proposed by previous scientists and developers was done. Based on the findings, the objectives of the system were fixed. As the power and advantages of AI-powered applications are trending, the project will be a state-of-the-art trending technology utilization. In this system, an overview is provided of how music can affect the user's mood and how to choose the right music tracks to improve the user's moods.

The implemented system can detect the user's emotions. The emotions that the system can detect were happy, sad, angry, neutral, or surprised. We integrate computer vision and machine learning techniques for connecting facial emotion for music recommendation. The approach is to use Deep Neural Networks (DNN) to learn the most appropriate feature abstractions. DNNs have been a recent successful approach in visual object recognition, human pose estimation, facial verification and many more. Convolution Neural Networks (CNNs) are proven to be very effective in are as such as image recognition and classification. The proposed system can detect facial expressions of the user using a CNN model. After determining the user's emotion, the proposed system provided the user with a playlist that contains music matches that detected the mood. Processing a huge dataset is memory as well as CPU intensive. This will make development more challenging and attractive. The motive is to create this application in the cheapest possible way and also to create it under a standardized device. The music recommendation system based on facial emotion recognition will reduce the efforts of users in creating and managing playlists. In this project, a main web page is designed where an image of the user is clicked. The image is then sent to the server to make the prediction about the emotion of the user. Once the emotion is detected, the next phase is to play songs. The songs from the database is displayed on the website.

The outcomes showed admirable performance, with high rates of emotion identification accuracy across different expressions. The system's capacity to correctly link corresponding musical playlists to emotional states supports the efficacy of the strategy. The GUI interface further improves user engagement and accessibility by offering a simple framework for natural communication.

However, it's critical to recognize some restrictions. Based on individual variations in face expression, the system's performance can change, and it might have trouble identifying more subtly expressed emotional states. Additionally, by adding more characteristics or using a multimodal method that combines voice analysis with face clues, the model's performance might be improved still further. Another possible area for expansion is the song dataset's relatively small size, which would enable a wider and more varied musical universe.

The study establishes a solid platform for advancements and new ideas in emotion-based music recommendation systems in the future. The combination of machine learning and emotional analysis has enormous potential, not just for music but also for more general applications like tailored content delivery across different industries. We foresee a more sophisticated and nuanced system that resonates more deeply with consumers, giving an enhanced music listening experience matched to their emotional inclinations, by continuously improving the model, expanding datasets, and investigating fresh methodologies. Thus, this initiative is a crucial step toward a day when technology and human feeling coexist harmoniously, altering how we interact with the musical world.

CHAPTER 10
BIBLIOGRAPHY

[1] Verma, R., & Bajaj, V. (2020). Emotion recognition through facial expression analysis using machine learning. International Journal of Advanced Research in Computer Science, 11(6), 1-6.

[2] Chen, X., Yang, X., & Jia, C. (2019). A Deep Learning Based Emotion Recognition System with Facial Expressions and Speech Signals. IEEE Access, 7, 53051-53063.

[3] Seppänen, T., & Väyrynen, E. (2019). Emotion detection and music recommendation using facial expression analysis. In Proceedings of the 27th European Signal Processing Conference (EUSIPCO), A Coruña, Spain, September 2-6, 2019.

[4] Singh, A., & Dangayach, G. S. (2021). A Machine Learning Approach for Emotion Detection through Facial Expression Recognition. In Proceedings of the 4th International Conference on Computer Science and Information Technology (CoSIT 2021), Cham, Switzerland, April 22-23, 2021.

[5] Kaggle Facial Expression Recognition dataset: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data.

[6] Rahul ravi, S.V Yadhukrishna, Rajalakshmi, Prithviraj, "A Face Expression Recognition Using CNN &LBP",2020 IEEE.

[7] Poria, S., Cambria, E., & Hussain, A. (2017). Context-aware multimodal emotion recognition using multimodal deep learning. Cognitive Computation, 9(3), 335-345.

[8] Liu, P., Qian, Y., Zhou, G., & Tang, X. (2017). Facial expression recognition: A brief tutorial overview. IEEE Signal Processing Magazine, 34(1), 35-52.

[9] Barsoum, E., Zhang, C., Canton Ferrer, C., & Zhang, Z. (2016). Training deep networks for facial expression recognition with crowd-sourced label distribution. In Proceedings of the 18th ACM International Conference on Multimodal Interaction (pp. 279-283).

[10] Gao, J., Zhou, Z., Zhang, D., & Li, M. (2018). Facial expression recognition based on deep evolutional spatial-temporal networks. IEEE Transactions on Image Processing, 27(7), 3434-3445.

[11] Li, X., Song, D., Zhao, T., & Fang, W. (2017). A survey of human computation techniques: Classification and applications. ACM Computing Surveys (CSUR), 50(2), 1-36.

[12] Picard, R. W. (1997). Affective computing. MIT press.

[13] Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2), 124-129.

[14] Lo, A. W. (2015). Big data and the complexity of industry classifications. Journal of Financial Data Science, 1(1), 2-12.

[15] Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161-1178.

[16] Sharma, A., & Paliwal, K. K. (2018). Speech emotion recognition: A review. Neural Computing and Applications, 29(3), 569-582.

EXECUTIVE SUMMARY

TITLE: Song recommendation based on User's Mood using Machine Learning

OBJECTIVES:

These are the objectives of the project:

1. The main objective of this project is to develop a personalized music recommendation system that can capture and analyze the user's mood with facial expression using machine learning algorithms, ultimately leading to a more engaging and satisfying music listening experience for the user.
2. Technology has led to the development of advanced music players with features like fast forward, reverse, variable playback speed, local playback, streaming playback with multicast streams, volume control, and genre classification. In this project the songs are classified based on users moods like sad, happy, etc.,
3. While these features can fulfill basic user needs, the task of sifting through playlists to locate songs that align with the user's current mood remains a manual process. To simplify this process, extensive playlists containing a variety of songs are associated with specific user moods. Upon identifying the user's prevailing mood, the corresponding playlist is promptly presented.
4. By collecting diverse dataset of images with corresponding mood labels, then training and validating the machine learning model, the system can recommend songs. If the person has a negative emotion, then a certain playlist will be shown that includes the most related types of music that will enhance his mood. And if the emotion is positive, a specific playlist will be presented which contains different types of music that will inflate the positive emotions.

RESEARCH METHODOLOGY:

The research methodology may include the following steps:

Collecting Datasets: The FER-2013 dataset from Kaggle which includes both training and testing datasets. The training set has 28,709 images and the testing set has 3,589 images, all of which are grayscale 48x48 pixel images of faces labeled with one of 7 emotions - happy, sad, angry, neutral, surprised, disgust and fear. The songs dataset from kaggle has 686 of songs ,about 100 songs per emotion, with its artist name and is classified into the 7 mood categories.

Face Detection and Feature Extraction: Face detection uses such classifiers, which are algorithms that detect what's either a face (1) or not a face (0) in an image. Classifiers are trained to detect faces using numbers of images to get more accuracy.he main aim of face detection is to spot the face within the frame by reducing external noises and other factors.During feature extraction, a pre- trained neural network is used as an arbitrary feature extractor.

Emotion Detection: The convolution neural network (CNN) architecture applies filters or feature detectors to an input image. The filters help identify various features in the image, such as edges, lines, and bends. By this the face is identified as either happy, sad, angry, neutral, surprised, disgust or fear.

Music Recommendation: The emotion module is used to detect real-time emotions of the user and labels them as happy, sad, angry, surprise, disgust, fear or neutral. These labels are connected to the songs database, then the program fetches the correct playlist according to the detected emotion and displays it in the GUI of the music player, with captions according to the detected emotions.

MAJOR FINDINGS:

In this system, we provide an overview of how music can affect the user's mood and how to choose the right music tracks to improve the user's moods. The motive was to create this application in the cheapest possible way and also to create it under a standardized device. The implementation of a CNN for facial emotion detection proved highly effective, it improved the efficiency of the emotion detection accuracy. The model's overall accuracy on the test dataset was 95%, demonstrating its effectiveness in identifying minute variations in facial features linked to various emotional states. The successful mapping of these emotions to corresponding music playlists underscores the viability of utilizing facial cues as a means to capture a user's mood, ultimately leading to more engaging and satisfying music listening experiences.

Processing a huge dataset is memory as well as CPU intensive. This made development more challenging and attractive. However, the project also highlighted the importance of diverse and inclusive datasets for training robust emotion detection models, emphasizing the need to encompass a broader spectrum of demographic and cultural expressions. Furthermore, challenges emerged in real-world conditions, with factors like varying lighting and occlusions impacting the system's performance. This music recommendation system based on facial emotion recognition will reduce the efforts of users in creating and managing playlists.

SUGGESTIONS:

First off, it would be advantageous to add more different face expressions to the dataset that represent a range of ages and cultural backgrounds. This would improve the model's capacity to recognize and react to a larger range of emotions. The accuracy and usability of the system could also be improved by investigating alternate emotion recognition techniques, such as multimodal approaches combining facial expressions with voice analysis. Additionally, taking into account real-time user preferences and feedback when modifying recommendations would result in a more dynamic and tailored user experience. Last but not least, regular updates and upkeep of the song database to reflect changing musical preferences and genres would guarantee the system's long-term relevance and user-engagement. Together, these ideas seek to improve the system's efficiency and customer pleasure while promoting a more thorough and personalized music recommendation experience.