

**A
Project Report
on**

**SONG RECOMMENDATION SYSTEM BASED ON
EMOTIONS**

Submitted to

Sant Gadge Baba Amravati University, Amravati

**Submitted in partial fulfilment of
the requirements for the Degree of
Bachelor of Engineering in**

Computer Science and Engineering

Submitted by

Virendra Janrao Pawar
(PRN: 213120431)

Sandeep Sanjay Daberao
(PRN: 213120221)

Sahil Pankaj Tawari
(PRN: 213120301)

Saad Khan Abutalib Khan
(PRN: 213120337)

Sankalp Sanjay Jumale
(PRN: 213120363)

**Under the Guidance of
Ms. B. S. Nawale
Assistant Professor CSE**



**Department of Computer Science and Engineering
Shri Sant Gajanan Maharaj College of Engineering,
Shegaon – 444 203 (M.S.)**

Session 2024-2025


**SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,
SHEGAON – 444 203 (M.S.)**


DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING




CERTIFICATE

This is to certify that **Mr. Virendra Pawar, Mr. Sahil Tawari, Mr. Saad Khan, Mr. Sandeep Daberao and Mr. Sankalp Jumale** students of final year Bachelor of Engineering in the academic year 2024-25 of Computer Science and Engineering Department of this institute have completed the project work entitled “**Song Recommendation System Based on Emotions**” and submitted a satisfactory work in this report. Hence recommended for the partial fulfilment of degree of Bachelor of Engineering in Computer Science and Engineering.


Ms. B. S. Nawale
Project Guide


Dr. J. M. Patil
Head of Department


Dr. S. B. Somani
Principal
SSGMCE, Shegaon

**SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,
SHEGAON – 444 203 (M.S.)**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

This is to certify that **Mr. Virendra Pawar, Mr. Sahil Tawari, Mr. Saad Khan, Mr. Sandeep Daberao and Mr. Sankalp Jumale** students of final year Bachelor of Engineering in the academic year 2024-25 of Computer Science and Engineering Department of this institute have completed the project work entitled “**Song Recommendation System Based on Emotions**” and submitted a satisfactory work in this report. Hence recommended for the partial fulfillment of degree of Bachelor of Engineering in Computer Science and Engineering.

Internal Examiner

Name and Signature

Date:

External Examiner

AR Ladole.
Name and Signature

Date:

Acknowledgement

It is our utmost duty and desire to express gratitude to various people who have rendered valuable guidance during our project work. We would have never succeeded in completing our task without the cooperation, encouragement and help provided to us by them. There are a number of people who deserve recognition for their unwavering support and guidance throughout this report.

We are highly indebted to our guide **Ms. B. S. Nawale** for her guidance and constant supervision as well as for providing necessary information from time to time. We would like to take this opportunity to express our sincere thanks, for her esteemed guidance and encouragement. Her suggestions broaden our vision and guided us to succeed in this work.

We are sincerely thankful to **Dr. J. M. Patil** (HOD, CSE Department, SSGMCE, Shegaon), and to **Dr. S B Somani** (Principal, SSGMCE, Shegaon) who always has been kind to extend their support and help whenever needed.

We would like to thank all teaching and non-teaching staff of the department for their cooperation and help. Our deepest thank to our parents and friends who have consistently assisted us towards successful completion of our work.

Virendra Pawar (22BD310804)
Sandeep Daberao (22BD310788)
Sahil Tawari (22BD310784)
Saad Khan (22BD310783)
Sankalp Jumale (22BD310790)

Contents

Abstract	i
List of Abbreviations	ii
List of Figures	iii
List of Screenshots	iv
List of Tables	v
1. Introduction	1
1.1 Overview	2
1.2 Motivation	4
1.3 Problem Statement	5
1.4 Objectives	5
1.5 Scope of Project	5
1.6 Organization of Project	6
2. Literature Review	8
2.1 Introduction to Proposed System	9
3. Methodology	13
3.1 System Architecture	14
3.2 Dataset Details	16
3.3 Algorithm Details	18

4. Implementation	21
4.1 Existing System	22
4.2 Proposed System	22
5. Result and Discussion	26
5.1 Libraries Used	27
5.2 Visualization Result	27
5.3 Result	30
6. Conclusion	35
6.1 Conclusion	36
6.2 Future Scope	36
References	37
Dissemination of Work	40
Plagiarism Report	51
Project Group Members	52

Abstract

This project presents an innovative system for song recommendation using a chatbot integrated with facial expression analysis and sentiment analysis. By leveraging generative AI, the Gemini API, and the Flask framework, the system personalizes song recommendations based on user input and detected emotional states. DeepFace and Convolutional Neural Networks (CNNs) are utilized for facial expression recognition, identifying emotions such as happy, surprised, sad, angry, disgust, and fear to classify user sentiment as positive, negative, or neutral. The dataset for song recommendations consists of 20th and 21st-century songs, categorized based on emotional tone to match user moods accurately. Additionally, a Kaggle-sourced facial expression dataset is used to train the CNN model for precise emotion detection. This combination of deep learning, emotion recognition, and AI-driven recommendations enhances user experience in digital entertainment, demonstrating the potential of AI in creating emotion-aware applications for personalized and immersive interactions.

List of Abbreviations

Abbreviation	Description
CNN	Convolution Neural Network
API	Application Programming Interface
AI	Artificial Intelligence
NLP	Natural Language Processing
VGG	Visual Geometry Group
RNN	Recurrent Neural Network

List of Figures

Figure No.	Description
Figure 3.1	flow chart of the project
Figure 3.4	The DeepFace library supports 7 state-of-the-art face recognition models.
Figure 5.2	Model Accuracy Graph

List of Screenshot

Screenshot No.	Description
Screenshot 3.2	Songs List Dataset
Screenshot 3.3	Dataset for Facial Expression
Screenshot 4.1	dataset image
Screenshot 5.3	Epochs details
Screenshot 5.4	Home Page of the project
Screenshot 5.5	Register page of the project
Screenshot 5.6	Login page of the project
Screenshot 5.7	Welcome of the project
Screenshot 5.8	Chatbot Option of the System
Screenshot 5.9	Chatbot Feature output of the System
Screenshot 5.10	Chatbot Facial Expression Detection
Screenshot 5.10	Chatbot Facial Expression output

CHAPTER 1
INTRODUCTION

INTRODUCTION

1.1 OVERVIEW

Music has always been a powerful medium for emotional expression and connection. With advancements in artificial intelligence, personalized recommendations have become increasingly sophisticated. This project explores the integration of facial expression analysis and sentiment detection into a song recommendation system, offering an engaging and emotionally aware user experience. By utilizing generative AI and APIs like Gemini, the system provides real-time, contextually relevant song suggestions. The incorporation of the Flask framework ensures a user-friendly and interactive platform for deploying the solution. This initiative aims to enhance the personalization and emotional resonance of digital music platforms.

A. Understanding Emotion Facial Detection

This essential component of computer vision aims to recognize and understand facial expressions to represent human emotions. This section provides a thorough analysis of the evolution of emotion face detection technologies throughout history, highlighting key discoveries, ethical issues, and underlying mechanics. We explore the inner workings of state-of-the-art facial recognition systems, illuminating both their strengths and weaknesses [\[1\]](#).

Evolution of face Detection Technology: Facial detection technology has come a long way, from simple face recognition systems to sophisticated emotive facial identification algorithms of today. Examining this route means examining the major turning moments, challenges encountered, and inventions that have brought technology thus far. The on-going endeavour to increase the accuracy and usefulness of face recognition systems is exemplified by the development of intricate algorithms and early attempts to recognize facial characteristics [\[2\]](#).

Mechanisms and Algorithms: Facial emotion recognition relies on a complex system of algorithms and mechanisms. The course progresses from an introduction to the fundamentals of facial feature extraction to an examination of the revolutionizing deep learning techniques. Engaging in this process will facilitate a deeper understanding of the underlying mechanisms that govern the conversion of facial expressions into emotions.

B. The Ability to recognize Emotions:

Emotion recognition, a critical component of the emotional facial detection procedure, is concerned with comprehending the emotional tone of facial expressions. The psychological foundations of human emotive identification and interpretation are explored in this section. Furthermore, an investigation is conducted into the function of machine learning in optimisation algorithms, which enables emotion recognition systems to attain an unparalleled level of accuracy and flexibility [3].

C. Chatbot Assistance: Bridging between Humans and Computers

Through facilitating communication between users and emotionally sophisticated algorithms, chatbots significantly impact the domain of human-computer interaction. This segment delves into the potential integration of chatbots and emotion face detection and identification technologies, examining their functionality and tracing their evolutionary paths [4].

1. Chatbot Technology's Development Over Time: The evolutionary progression of chatbots from basic rule-based systems to advanced integrations of machine learning and natural language processing (NLP) in the twenty-first century shows their revolutionary development. The article investigates the wide-ranging implementation of chatbot integration across various industries, encompassing its utilization as personal assistant and its potential to revolutionize customer service interactions. This conversation aims to demonstrate how advancements in technology have empowered chatbots to surmount the constraints of predetermined responses and engage in more dynamic and contextually relevant exchanges with users.

2. Machine Learning and Natural Language Processing: This segment explores the technical foundations of natural language processing (NLP) and machine learning, underscoring their critical importance in the operation of chatbots.. The analysis reveals the dynamic nature of these conversational agents by investigating how these technologies empower chatbots to understand user queries, participate in nuanced and contextually appropriate dialogues, and consistently improve their performance. Furthermore, by combining NLP and machine learning, chatbots are enabled to evolve and conform to the requirements and expectations of users, thereby promoting more natural and fluid interactions.

3. Emotional Interfaces with Chatbots: With respect to chatbots functioning as emotive interfaces, this subsection explores unexplored domains. While their main function is to relay information, researchers are currently examining whether chatbots have the ability to detect and respond to user emotions with the intention of delivering a more personalized and empathetic user experience. In this section, the consequences for user engagement, satisfaction, and the overall caliber of human-computer interactions are assessed as the mechanisms underlying chatbots' ability to derive emotional signals are explored. By serving as emotional interfaces, chatbots enable the development of a digital environment that is more intuitive and responsive; within this environment, user emotions are not only acknowledged but also appropriately addressed[5].

D. Music as a Catalyst of Emotions:

Music, widely acknowledged as a potent conduit for evoking and articulating emotions, occupies a central position in the human condition. This segment examines the profound affective influence of music, investigates the psychological foundations of musical inclinations, and scrutinizes the revolutionary possibilities that arise from the integration of facial recognition and emotion detection technologies in order to construct playlists customized to individual dispositions.

1.2 MOTIVATION

In today's fast-paced world, music plays a vital role in influencing emotions, reducing stress, and enhancing productivity. However, traditional song recommendation systems primarily rely on user history, preferences, and genre-based filtering, often neglecting real-time emotional states. This project aims to bridge that gap by integrating facial expression analysis and sentiment analysis into a song recommendation system. With the rise of AI-driven conversational agents, chatbots have become an essential medium for human-computer interaction. By incorporating generative AI and the Gemini API, this project enhances chatbot capabilities, making song recommendations dynamic, emotion-aware, and highly personalized[6].

Furthermore, with the growing interest in emotional AI, this project explores the potential of computer vision and natural language processing (NLP) to understand human emotions more effectively. The combination of these technologies creates an interactive experience where the system can detect emotions in real-time and suggest

appropriate music, catering to users' moods instantly. By leveraging AI and the Flask framework, this project aims to enhance user engagement in digital entertainment while demonstrating the potential of emotional intelligence in AI-driven applications.

1.3 PROBLEM STATEMENT

Despite the abundance of music streaming platforms, existing systems often lack the ability to adapt to users' emotional states in real-time. Traditional recommendation algorithms focus on historical data and fail to account for the dynamic nature of human emotions. This project addresses this gap by introducing a system that combines sentiment analysis and facial expression recognition to provide emotionally intelligent song recommendations.

1.4 OBJECTIVES

- To develop a chatbot-based song recommendation system that uses facial expression analysis and sentiment detection to enhance personalization.
- To utilize generative AI and APIs like Gemini to optimize the recommendation process.
- To create a seamless and engaging user interface using the Flask framework.

1.5 SCOPE OF PROJECT

The system utilizes real-time emotion recognition through computer vision techniques to classify facial expressions as happy, surprised, or sad, categorizing them into positive, negative, or neutral sentiments. Additionally, it employs NLP-based sentiment analysis to extract emotional cues from user text input, refining the recommendation process. By integrating generative AI via the Gemini API, the chatbot interactively enhances recommendations based on user queries and preferences. A curated dataset of songs ensures personalized suggestions that align with the detected emotional state. Built using the Flask framework, the system is accessible as a real-time web application, offering an engaging and interactive experience. By combining facial recognition, sentiment analysis, and AI-powered conversations, this project aims to set new benchmarks for AI-driven recommendation systems in the digital entertainment industry[7].

1.6 ORGANIZATION OF PROJECT

This thesis is structured into seven main chapters, each focusing on different aspects of the research, from the introduction to the final conclusions and future work. Below is an overview of the organization:

1. Introduction

This chapter provides an overview of the project, including the background, motivation, problem statement, objectives, scope, and limitations. It highlights the significance of AI-driven song recommendations using facial expression and sentiment analysis.

2. Literature Review

This chapter explores previous research and existing methodologies related to emotion detection, music recommendation systems, and deep learning-based sentiment analysis. It compares various approaches and identifies gaps that this project aims to address.

3. Research Methodology

This section details the system architecture, dataset selection, preprocessing techniques, and algorithm implementation. It describes the working of DeepFace, CNN models, and the Gemini API used for generating personalized song recommendations.

4. Implementation

This chapter discusses the existing system and its limitations, followed by the proposed system and its advantages. It includes a step-by-step explanation of how the system processes facial expressions and recommends songs based on emotions.

5. Deployment and Results

This chapter explains the system requirements, visualization of results, deployment process, and final outcomes. It presents the performance metrics of the CNN model, DeepFace accuracy (97.3%), and overall system efficiency, validating its effectiveness.

6. Conclusion

This chapter summarizes the key findings of the project, emphasizing the

success of integrating AI and deep learning techniques in emotion-aware music recommendations. It highlights the CNN model's accuracy (70.92% training and 63.92% validation) and the impact of generative AI in enhancing digital entertainment.

7. Future Scope

This section discusses potential improvements and advancements for the project, such as expanding the dataset, integrating real-time emotion detection, supporting multilingual song recommendations, and enhancing system scalability for a larger user base.

CHAPTER 2
LITERATURE
REVIEW

LITERATURE REVIEW

2.1 INTRODUCTION TO PROPOSED SYSTEM

Below, we discuss the literature review of the project, highlighting existing research on AI-driven recommendation systems, emotion recognition, and chatbot-based interactions. This review provides insights into the advancements and challenges in integrating facial expression analysis and sentiment-based recommendations.

1. Regin, R., S. Suman Rajesh, and T. Shynu, The paper highlights that the automated conversation system developed using Natural Language Processing (NLP) in Python effectively handles real-time user queries with high accuracy and contextual understanding. Leveraging Python's extensive libraries and machine learning models, the chatbot demonstrated scalability, multilingual support, and adaptability for domain-specific applications such as customer support, education, and healthcare. The study revealed that combining rule-based and learning-based approaches enhanced response accuracy and user interaction quality. Additionally, the implementation was found to be cost-effective and showcased significant potential for future advancements, including integration with voice recognition and advanced sentiment analysis[1].
2. Aleedy, Moneerh, The paper explores the generation and analysis of chatbot responses using Natural Language Processing (NLP) techniques to enhance conversational quality and user engagement. It demonstrates the effectiveness of combining rule-based methods and machine learning algorithms for generating contextually appropriate and coherent responses. The study emphasizes the importance of sentiment analysis and intent recognition in improving the chatbot's ability to understand user queries and provide relevant answers. Furthermore, it highlights the role of evaluation metrics in assessing response quality and refining chatbot models. The findings underscore the potential of NLP-powered chatbots in delivering dynamic and meaningful interactions across various domains[2].
3. Chamishka, Sadil, et al, The paper presents a voice-based real-time emotion detection technique that utilizes Recurrent Neural Networks (RNN) for

advanced feature modeling. By analyzing vocal attributes such as pitch, tone, and intensity, the system effectively identifies emotional states with high accuracy and responsiveness. The study highlights the superiority of RNNs in capturing temporal dependencies in audio signals, making them well-suited for real-time emotion recognition tasks. The proposed technique demonstrates robust performance across diverse datasets and scenarios, underscoring its applicability in fields like mental health monitoring, human-computer interaction, and customer service. The findings emphasize the system's potential to enhance user experiences through real-time emotional insights[3].

4. Kumar et al. The paper provides a comprehensive review of chatbot design and implementation techniques, exploring both rule-based and artificial intelligence-driven approaches. It highlights the evolution of chatbot systems, focusing on advancements in Natural Language Processing (NLP), machine learning, and conversational design principles. The study identifies key challenges such as maintaining context, handling ambiguous inputs, and achieving high user satisfaction. It emphasizes the importance of integrating sentiment analysis, intent recognition, and adaptive learning mechanisms for creating effective and user-centric chatbots. Additionally, the review underscores the potential of chatbots in various domains, including healthcare, customer service, and education, while advocating for further research in enhancing scalability, personalization, and multimodal interaction capabilities[4].
5. Alu,et al. The paper investigates the use of Convolutional Neural Networks (CNNs) for voice-based emotion recognition in companion robots, focusing on enhancing human-robot interactions. By analyzing audio features such as mel-frequency cepstral coefficients (MFCCs) and spectrograms, the proposed system effectively identifies emotional states with high accuracy. The study demonstrates the robustness of CNNs in extracting complex patterns from voice data, making them well-suited for real-time applications in dynamic environments. The findings highlight the system's potential to improve the emotional intelligence of companion robots, enabling them to respond empathetically and adaptively to user emotions, thereby enhancing user

experience and acceptance in domains like elder care and mental health support[5].

6. Touimi, The paper introduces an intelligent chatbot integrated with a Latent Dirichlet Allocation (LDA)-based recommender system to enhance user interactions and provide personalized recommendations. By leveraging LDA for topic modeling, the chatbot identifies user preferences and delivers relevant suggestions in real time. The study demonstrates the system's capability to handle diverse queries while maintaining contextual coherence, thereby improving user satisfaction. The integration of the chatbot with the LDA framework enables dynamic adaptability across various domains, such as education and e-commerce. The findings emphasize the effectiveness of combining NLP techniques with advanced recommendation algorithms for delivering intelligent, user-centric solutions[6].
7. Smith, et al The paper provides a detailed review of emotion analysis techniques in music recommender systems, emphasizing their role in enhancing personalization and user engagement. It highlights advancements in emotion recognition methods, including audio signal analysis, lyric sentiment analysis, and multimodal approaches that combine audio and text features. The study identifies the effectiveness of machine learning algorithms, such as deep learning models, in capturing emotional nuances in music. Additionally, it underscores the challenges of emotion variability across cultures and individual preferences. The findings suggest that integrating emotion analysis into recommender systems significantly improves user satisfaction by aligning music recommendations with users' emotional states and contexts[7].
8. Garcia, the paper investigates user satisfaction with emotion-aware music recommender systems, focusing on their ability to align music suggestions with users' emotional states. It highlights that incorporating emotion recognition techniques, such as sentiment analysis and audio feature extraction, significantly enhances the personalization and relevance of recommendations. The study finds that users perceive emotion-aware systems as more engaging and satisfying compared to traditional recommendation approaches. However, it also identifies challenges, such as accurately

detecting user emotions in real time and addressing individual variations in emotional responses to music. The findings underscore the potential of emotion-aware recommender systems to transform user experiences in music consumption by fostering deeper emotional connect[\[8\]](#).

CHAPTER 3
METHODOLOGY

METHODOLOGY

3.1 SYSTEM ARCHITECTURE

The AI-based music recommendation system illustrated in Fig 3.1 represents the overall flow of how user interactions are handled to deliver personalized song suggestions. The process begins with user interaction, where new users register and existing users log in. After successful authentication, users are directed to a main dashboard where they interact with an AI-powered chatbot. This chatbot plays a crucial role in detecting user emotions through two primary methods: facial emotion recognition and text-based sentiment analysis.

Facial recognition leverages a deep learning model that analyzes real-time facial expressions using the device's camera. It examines features such as eye movement, mouth shape, and facial muscle activity to identify emotional states like happy, sad, angry, or neutral. Simultaneously, the system processes the user's textual input using natural language processing and sentiment analysis to determine the underlying emotion in the text.

Once the emotion is identified through either method, the system maps the detected emotion to an appropriate category of music. For instance, happy emotions trigger recommendations of energetic and upbeat songs, while sad emotions prompt slow and soothing tracks. Neutral or calm moods may lead to soft instrumental or ambient music. The recommendation engine then fetches songs from a pre-defined dataset, where each song is tagged with mood-related metadata, ensuring the selection aligns with the user's current emotional state.

This entire architecture is designed to deliver a dynamic and personalized music experience by combining facial recognition, AI-based text analysis, and intelligent music recommendation in real time.

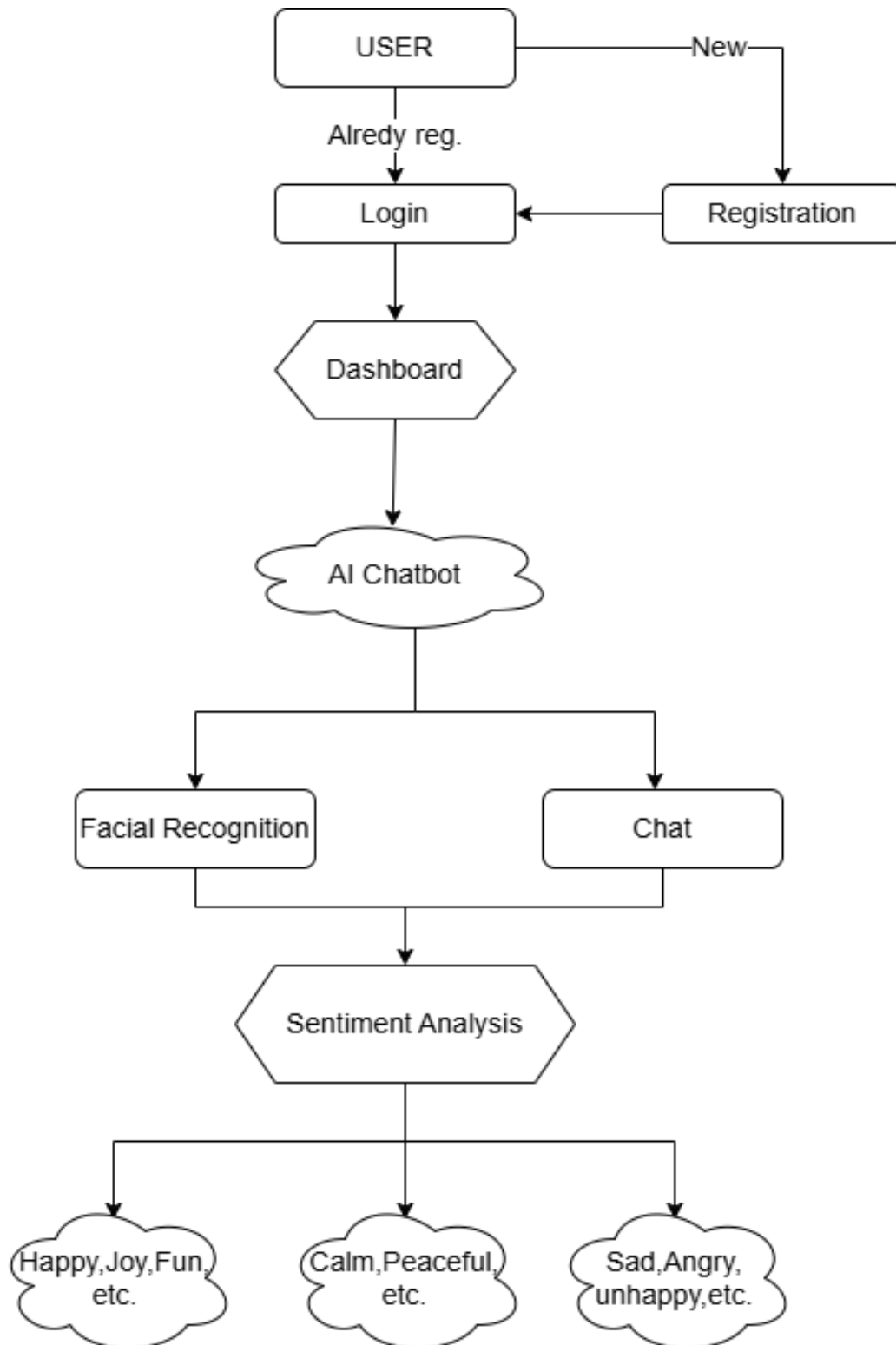


Fig 3.1 shows the flow chart of the project

3.2 DATASET DETAILS

The dataset used in this project consists of a curated collection of latest Bollywood songs, categorized based on three primary emotional tones: happy, sad, and neutral. Happy songs include energetic and uplifting tracks, sad songs feature slow and melancholic melodies, while neutral songs have a balanced or relaxing tone. The classification is based on lyrical themes, acoustic properties, and historical audience preferences to enhance AI-driven recommendations. Additionally, for facial expression recognition, a dataset from the Kaggle website has been utilized, which contains diverse human facial expressions labeled as happy, sad, surprised, neutral, and other emotional states.

 aaj_ki_raat	Aaj Ki Raat(PagalWorld.co...	Madhubanti Bagc...	Stree 2 (2024) - Pagal...
 Aankh_Hai_Bhari-Bh...	Aankh Hai Bhari Bhari (Du...	Kumar Sanu; Alka ...	Aankh Hai Bhari Bhari...
 Aap_Ke_Aa_Jane_Se	Aap Ke Aa Jane Se	Mohammed Aziz; ...	Khudgarz (Original M...
 Aaye_Ho_Meri_Zind...	Aaye Ho Meri Zindagi Me...	Udit Narayan	Raja Hindustani (Orig...
 Ae_Kash_Ke_Hum	Ae Kash Ke Hum	Jatin-Lalit; Kumar ...	Kabhi Haan Kabhi Na...
 Aisi_Deewangi			
 Aye_Zindagi_Gale_L...	Aye Zindagi Gale Lagaa Le	Ilaiyaraaja; Suresh ...	Sadma (Original Moti...
 Barsaat_Ke_Mausa...	Barsaat Ke Mausam Mein	Kumar Sanu; Roop...	Naajayaz (Original M...
 Bin_Tere_Sanam	Bin Tere Sanam	Udit Narayan; Kavi...	Yaara Dildara (Origin...
 Chand_Chhupa_Bad...	Chand Chhupa Badal Mein	Udit Narayan; Alka...	Hum Dil De Chuke Sa...
 Channe_Ke_KhetMein	Channe Ke Khet Mein	Poomima	Anjaam (Original Mot...
 Chhu_Kar_Mere_Ma...	Chhu Kar Mere Manko - Fr...	Kishore Kumar	Yaarana (Original Mo...
 Chhupana_Bhi_Nahi...	Chhupana Bhi Nahin Aata	Vinod Rathod	Baazigar (Original M...
 Dil_Ka_Aalam	Dil Ka Aalam (From "Aash...	Kumar Sanu	Dheere Dheere Se - T...
 Dil_Teri_Deewangi_...	Dil Teri Deewangi Mein	Anand Raj Anand; ...	Kismet
 Gazab_Ka_Hai_Din	Gazab Ka Hai Din	Alka Yagnik; Udit ...	Qayamat Se Qayamat...
 Gore_Gore_Mukhde...	Gore Gore Mukhde Pe	Udit Narayan; Alka...	Suhaag (Original Mot...
 Heartbeat	Heartbeat - Instrumental	Shankar-Ehsaan-Loy	Kal Ho Naa Ho (Orig...
 Hum_Aapke_Dil_Me...	Hum Aapke Dil Mein Reht...	Anuradha Paudwal...	Ultimate Romantic D...
 Humko_Sirf_Tumse	Humko Sirf Tumse	Kumar Sanu; Alka ...	Barsaat (Original Mot...
 Husn_Hai_Suhana	Husn Hai Suhana	Chandana Dixit; A...	Coolie No.1 (Original ...
 Jimmy_Jimmy_Jimm...	Jimmy Jimmy Jimmy Aaja	Parvati Khan	Disco Dancer
 Jiya_Dhadak_Dhada...	Jiya Dhadak Dhadak Jaye	Rahat Fateh Ali Kh...	Kalyug
 Jiya_Jale	Jiya Jale (From "Dil Se")	Lata Mangeshkar; ...	Lata Mangeshkar - Th...
 Jooma_Chumma_D...	Jooma Chumma De De	Sudesh Bhosle; Ka...	Hum
 Kal_Ho_Naa_Ho	Kal Ho Naa Ho	Shankar-Ehsaan-L...	Kal Ho Naa Ho (Orig...
 Kehna_Hi_Kya	Kehna Hi Kya - From "Bo...	K. S. Chithra	Bombay (Original Mo...
 Koi_Mil_Gaya	Koi Mil Gaya	Jatin-Lalit; Kavita ...	Kuch Kuch Hota Hai (...
 Mars_Dil_Bhi_Kitna	Mars Dil Bhi Kitna Dagal	Kumar Sanu; Alka ...	Bollywood Music - Ki...

Fig 3.2 shows the Songs List Dataset



Fig 3.3 (a) shows the Dataset for Facial Expression

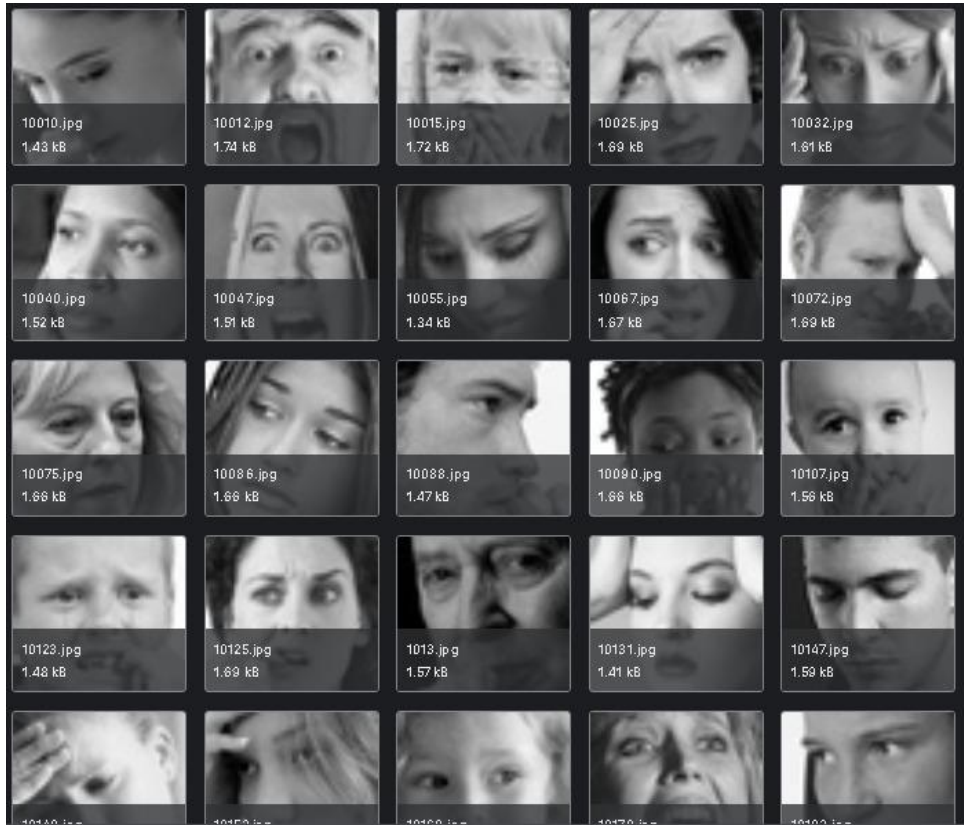


Fig 3.3 (b) shows the Dataset for Facial Expression

3.3 ALGORITHM DETAILS

CNN ALGORITHM

For the Recommendation System Using AI Chatbot, the Convolutional Neural Network (CNN) is primarily used for facial expression recognition to classify user emotions (*happy, sad, neutral, etc.*). The CNN architecture for project consists of the following layers:

CNN Architecture for Facial Expression Recognition:

1. Input Layer:
 - Takes an input image of the user's face (preprocessed grayscale or RGB image).
 - Resized to a fixed dimension (e.g., 48x48 or 64x64 pixels).
2. Convolutional Layers:
 - Multiple convolutional layers with 3x3 filters to extract facial features.
 - ReLU activation function is applied to introduce non-linearity.
3. Pooling Layers (MaxPooling):
 - 2x2 max-pooling layers to reduce spatial dimensions while retaining essential features.
4. Additional Convolutional & Pooling Layers:
 - Deeper feature extraction using additional Conv + ReLU + MaxPooling layers to capture complex facial expressions.
5. Flatten Layer:
 - Converts the extracted feature maps into a 1D vector for input into the dense layers.
6. Fully Connected Layers (Dense Layers):
 - A fully connected (FC) layer with 128 or 256 neurons to learn high-level representations.
 - Dropout layer (e.g., 0.5 dropout rate) to prevent overfitting.
7. Output Layer:
 - Softmax activation function for multi-class classification (*happy, sad, neutral, surprised, etc.*).

DEEPPFACE

DeepFace AI is the most lightweight face recognition and facial attribute analysis library for Python. The open-sourced DeepFace library includes all leading-[edge AI models](#) for modern face recognition and automatically handles all procedures for facial recognition in the background.

FEATURES

- **Face Verification:** The task of face verification refers to comparing a face with another to verify if it is a match or not. Hence, face verification is commonly used to compare a candidate's face to another. This can be used to confirm that a physical face matches the one in an ID document.
- **Face Recognition:** The task refers to finding a face in an image database. Performing face recognition requires running face verification many times.
- **Facial Attribute Analysis:** The task of facial attribute analysis refers to describing the visual properties of face images. Accordingly, facial attributes analysis is used to extract attributes such as age, gender classification, emotion analysis, or race/ethnicity prediction.
- **Real-Time Face Analysis:** This feature includes testing face recognition and facial attribute analysis with the real-time video feed of your webcam.

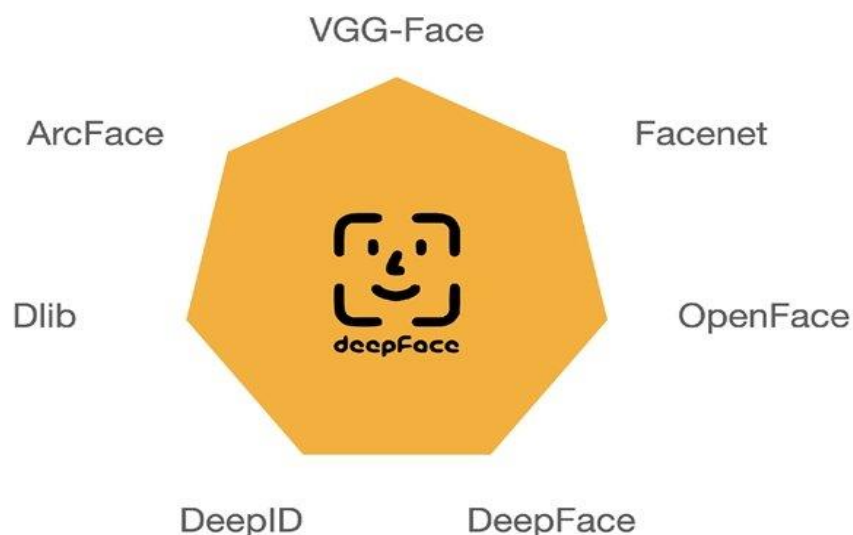


Fig 3.4 shows The DeepFace library supports 7 state-of-the-art face recognition models.

DeepFace has been expanding its model portfolio since its first commit. Its initial version wraps just VGG-Face and Facenet. It supports seven cutting-edge face recognition models. But also in the time to come, you will be able to easily use the latest face recognition models with DeepFace, because the model name is an argument of its functions, and the interface always stays the same.

CHAPTER 4
IMPLEMENTATION

IMPLEMENTATION

4.1 EXISTING SYSTEM

In the current music recommendation systems, most platforms rely on user preferences, history, and manual inputs to suggest songs. Traditional approaches use collaborative filtering or content-based filtering, where recommendations are made based on a user's past listening behavior or the attributes of songs they have previously liked. However, these methods lack real-time emotional adaptability, meaning they do not consider the user's instantaneous mood or facial expressions when providing recommendations. Additionally, text-based sentiment analysis is commonly used in chatbots to understand user emotions, but it is limited by the accuracy of natural language processing (NLP) and cannot capture non-verbal emotional cues such as facial expressions[9].

Some existing AI-driven chatbots integrate basic sentiment analysis to improve recommendations, but they do not leverage real-time facial recognition for emotion detection. Additionally, they often rely on predefined rule-based models rather than deep learning-based approaches such as Convolutional Neural Networks (CNNs) for facial expression analysis. Due to this limitation, they may fail to provide highly personalized and emotion-driven song recommendations. Moreover, most systems lack an interactive AI-powered refinement feature, where users can further customize recommendations based on real-time feedback. This gap in traditional systems highlights the need for an advanced AI-based chatbot that integrates computer vision, NLP-based sentiment analysis, and Generative AI to provide a more accurate and engaging music recommendation experience[10].

4.2 PROPOSED SYSTEM

To overcome the limitations of traditional music recommendation systems, this project introduces an AI-driven chatbot that integrates DeepFace for facial expression analysis, CNN for image classification, sentiment analysis, and generative AI to recommend songs based on real-time user emotions. The system follows a structured pipeline, including data collection, data preprocessing, feature selection, and model selection, to enhance recommendation accuracy and user engagement[11].

1) Data Collection

The dataset used in this project consists of two primary sources:

- **Facial Expression Dataset:** A publicly available dataset from Kaggle is used to train the DeepFace model for facial emotion recognition. This dataset contains labeled images of facial expressions such as happy, sad, neutral, surprised, angry, and fear to accurately map emotions.
- **Music Dataset:** A manually curated dataset of songs from the 2000s and 2024s is categorized based on emotional tone. The songs are labeled as happy, sad, or neutral to match the detected facial expressions and sentiment analysis results. This dataset is stored in a structured format for real-time retrieval and recommendation.



Fig 4.1 shows the dataset image

2) Data Preprocessing

Data preprocessing is an essential step to clean and standardize input data before feeding it into the models.

- **Facial Expression Data (DeepFace & CNN Processing):**
 - Images are resized to 48x48 or 64x64 pixels to ensure uniformity.
 - Grayscale conversion reduces computational complexity and enhances model efficiency.
 - DeepFace pre-trained models (VGG-Face, OpenFace, Facenet, DeepID) are used to extract facial embeddings.
 - CNN layers are applied to detect essential facial features for emotion classification.
 - Data augmentation (rotation, flipping, brightness adjustment) is implemented to increase dataset variability and prevent overfitting.

- **Text-Based Sentiment Analysis:**
 - User text inputs from the chatbot are tokenized and vectorized using NLP techniques such as TF-IDF or word embeddings (BERT, Word2Vec).
 - Stopwords, special characters, and redundant words are removed to eliminate noise.

- **Music Data Processing:**
 - Each song is tagged with metadata such as mood, genre, and tempo, which aids in matching songs to detected emotions.

3) Feature Selection

Feature selection ensures the system focuses on the most critical data points, improving accuracy and efficiency.

- **Facial Emotion Features:** DeepFace extracts 128-dimensional embeddings from faces, which are further processed by CNN layers to classify emotions.

- **Sentiment Analysis Features:** NLP models extract text-based sentiment scores using Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM) models.
- **Music Recommendation Features:**
 - Songs are analyzed based on tempo, pitch, mood, and lyrics sentiment to match them with facial and text-based emotions.
 - The chatbot also considers user history and preferences for better personalization.

4) Model Selection

The system integrates multiple AI models to enhance emotion recognition and music recommendations:

- **Facial Emotion Recognition:**
 - DeepFace is used for real-time facial emotion detection.
 - A custom CNN model is trained to improve accuracy by extracting local and global facial features.
- **Sentiment Analysis:** A Bi-LSTM (Bidirectional Long Short-Term Memory) model is used to analyze textual input from the chatbot and predict sentiment intensity.
- **Music Recommendation:**
 - A hybrid content-based and sentiment-driven recommendation model is implemented.
 - The Gemini API (Generative AI) enhances responses by suggesting songs dynamically based on user queries and detected emotions.

This proposed system enhances traditional recommendation approaches by leveraging DeepFace, CNN, and sentiment-driven AI models to provide a real-time, emotion-based, and highly personalized music recommendation experience.

CHAPTER 5
RESULT AND DISCUSSION

RESULT AND DISCUSSION

5.1 LIBRARIES USED

Deep Learning & Machine Learning

- TensorFlow – For building and training CNN models.
- Keras – High-level deep learning API used with TensorFlow.
- DeepFace – For facial emotion detection using pre-trained models.
- OpenCV – For image processing and facial feature extraction.
- scikit-learn – For data preprocessing, feature selection, and classification.
- Pandas & NumPy – For data handling and manipulation.

5.2 VISUALIZATION RESULT

This graph represents the model accuracy of Recommendations System Using AI chatbot project over 50 training epochs.

1. X-axis (Epochs): This represents the number of times your model has been trained on the entire dataset.
2. Y-axis (Accuracy): This represents how well your model is performing.
3. Train Accuracy (Blue Line): Shows how well the model is learning from the training data.
4. Validation Accuracy (Orange Line): Indicates how well the model generalizes to unseen data.

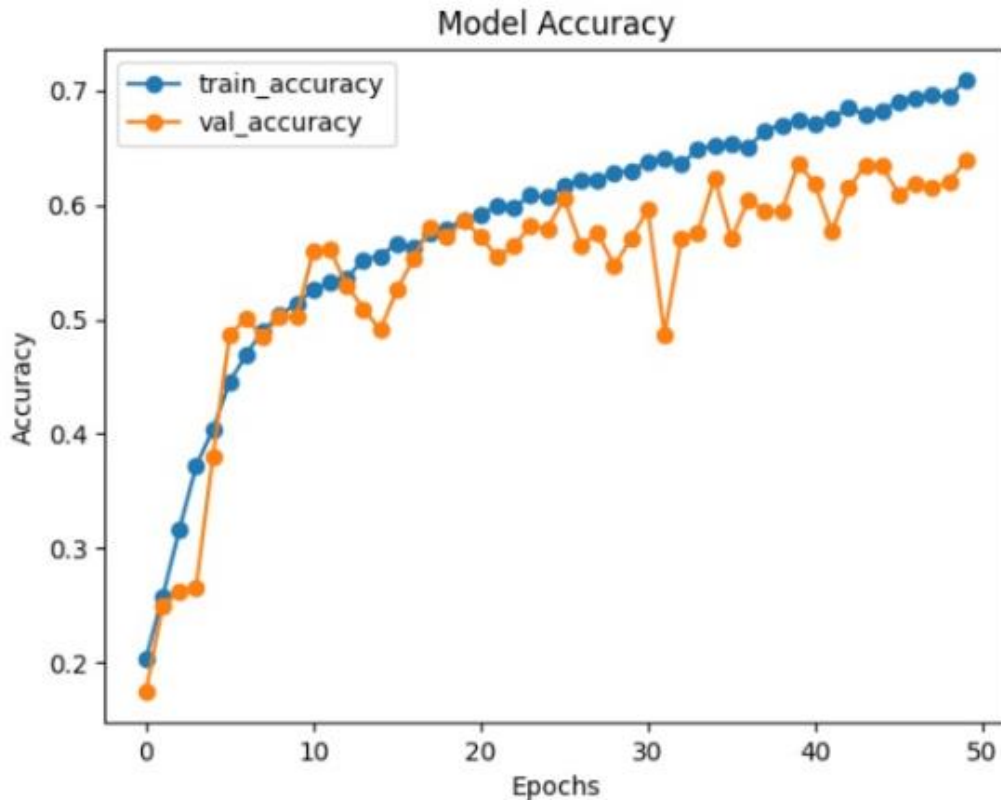


Fig 5.2 shows the Model Accuracy Graph.

1 Observations:

- Initial Phase (0-10 Epochs): Both training and validation accuracy are increasing steadily, indicating that the model is learning.
- Middle Phase (10-30 Epochs): Training accuracy continues to improve, while validation accuracy fluctuates but follows a similar trend.
- Later Phase (30-50 Epochs): The training accuracy keeps increasing, reaching around 0.7 (70%), while validation accuracy fluctuates more.

TRAINING RESULT

Training progress over 50 epochs, with key metrics:

- Loss (Training & Validation): Measures error. Lower is better.
- Accuracy (Training & Validation): Measures correctness. Higher is better.

2 Key Observations:

1. Initial Performance (Epoch 1)

- Training Loss: 2.2023
- Training Accuracy: 20.42%
- Validation Loss: 2.1801
- Validation Accuracy: 17.52%
- Interpretation: The model starts with random weights and learns patterns over time.

Interpretation: The model achieves good generalization, with training and validation accuracy increasing.

```
Epoch 1/50
WARNING:tensorflow:From c:\Users\gaiku\AppData\Local\Programs\Python\Python311\Lib\site-packages\keras\src\utils\tf_utils.py:492: The name tf.nn.conv2d is deprecated. Please use tf.nn.conv2d_v2 instead.
WARNING:tensorflow:From c:\Users\gaiku\AppData\Local\Programs\Python\Python311\Lib\site-packages\keras\src\engine\base_layer_utils.py:384: The name tf.executing_eager_session_on_cpu is deprecated. Please use tf.compat.v1.executing_eager_session_on_cpu instead.

163/163 [=====] - 247s 1s/step - loss: 2.2023 - accuracy: 0.2042 - val_loss: 2.1801 - val_accuracy: 0.1752
Epoch 2/50
163/163 [=====] - 240s 1s/step - loss: 1.8975 - accuracy: 0.2577 - val_loss: 2.8182 - val_accuracy: 0.2502
Epoch 3/50
163/163 [=====] - 222s 1s/step - loss: 1.7446 - accuracy: 0.3168 - val_loss: 2.3715 - val_accuracy: 0.2624
Epoch 4/50
163/163 [=====] - 220s 1s/step - loss: 1.6149 - accuracy: 0.3720 - val_loss: 2.1770 - val_accuracy: 0.2663
Epoch 5/50
163/163 [=====] - 226s 1s/step - loss: 1.5297 - accuracy: 0.4038 - val_loss: 1.6148 - val_accuracy: 0.3803
Epoch 6/50
163/163 [=====] - 246s 2s/step - loss: 1.4481 - accuracy: 0.4453 - val_loss: 1.3463 - val_accuracy: 0.4866
Epoch 7/50
163/163 [=====] - 247s 2s/step - loss: 1.3831 - accuracy: 0.4693 - val_loss: 1.2918 - val_accuracy: 0.5004
Epoch 8/50
163/163 [=====] - 247s 2s/step - loss: 1.3293 - accuracy: 0.4905 - val_loss: 1.3288 - val_accuracy: 0.4848
Epoch 9/50
163/163 [=====] - 245s 2s/step - loss: 1.2966 - accuracy: 0.5084 - val_loss: 1.3088 - val_accuracy: 0.5026
Epoch 10/50
163/163 [=====] - 245s 2s/step - loss: 1.2714 - accuracy: 0.5138 - val_loss: 1.2805 - val_accuracy: 0.5022
Epoch 11/50
...
Epoch 49/50
163/163 [=====] - 246s 2s/step - loss: 0.8149 - accuracy: 0.6944 - val_loss: 1.0450 - val_accuracy: 0.6197
Epoch 50/50
163/163 [=====] - 246s 2s/step - loss: 0.7917 - accuracy: 0.7092 - val_loss: 1.0543 - val_accuracy: 0.6392
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

Fig 5.3 shows the Epochs details

2. Mid Training Progress (Epoch 10)

- Training Loss: **1.2714**
- Training Accuracy: **51.38%**
- Validation Loss: **1.2805**
- Validation Accuracy: **50.22%**
- **Interpretation:** The model improves significantly, learning key features.

3. Final Performance (Epoch 50)

- Training Loss: **0.7917**
- Training Accuracy: **70.92%**
- Validation Loss: **1.0543**

- Validation Accuracy: **63.92%**

5.3 RESULT

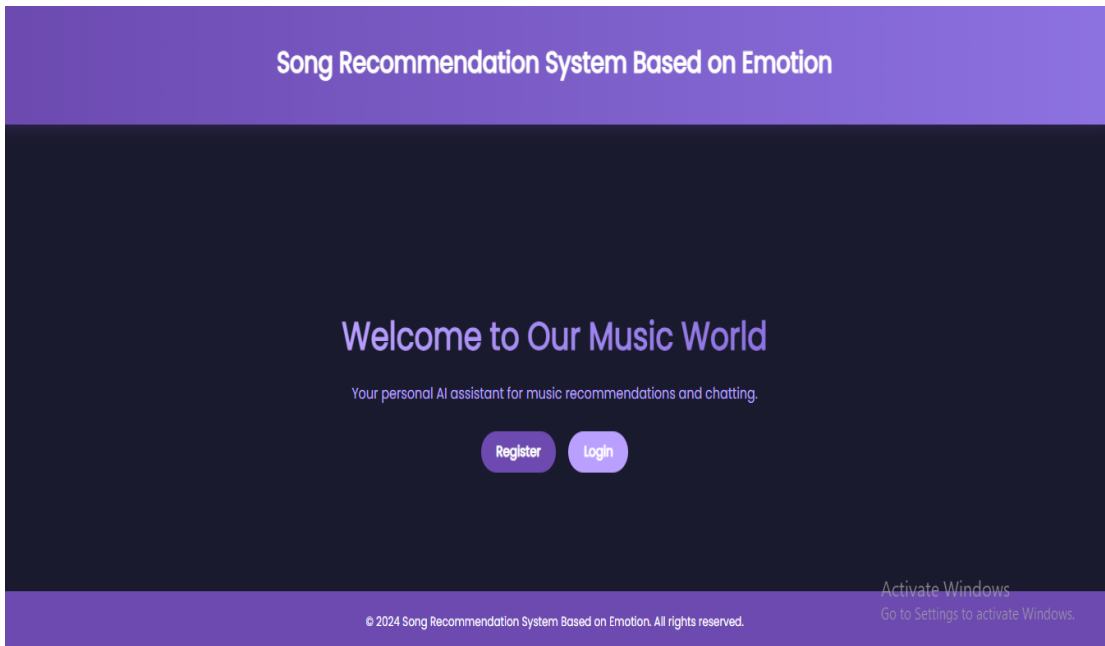


Fig 5.4 shows the Home Page of the project

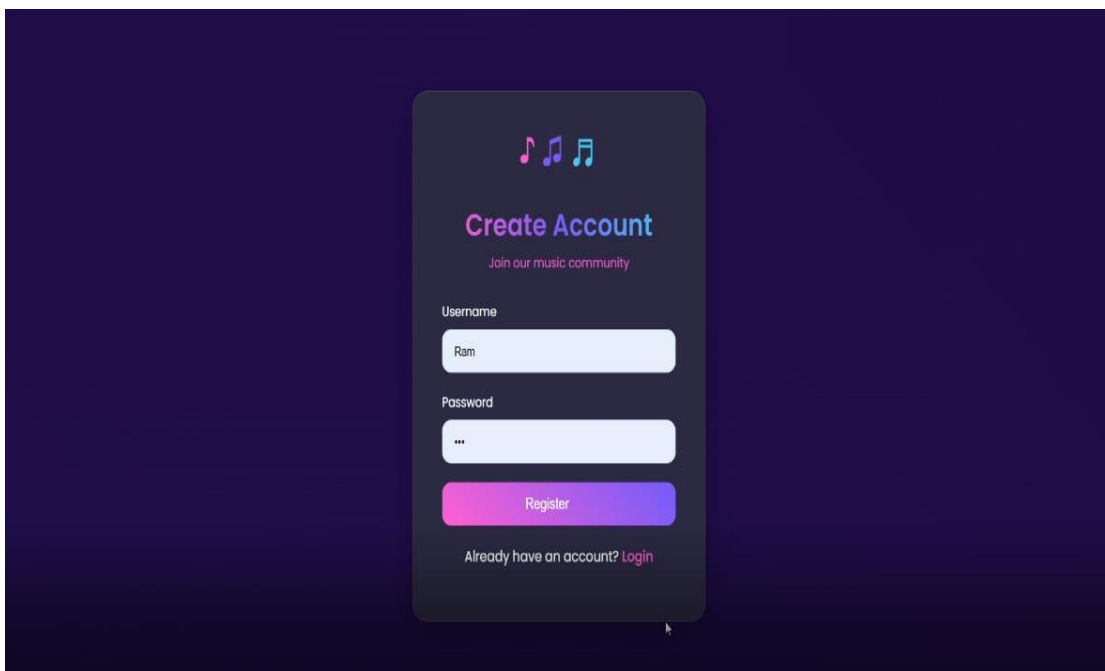


Fig 5.5 shows the Register page of the project

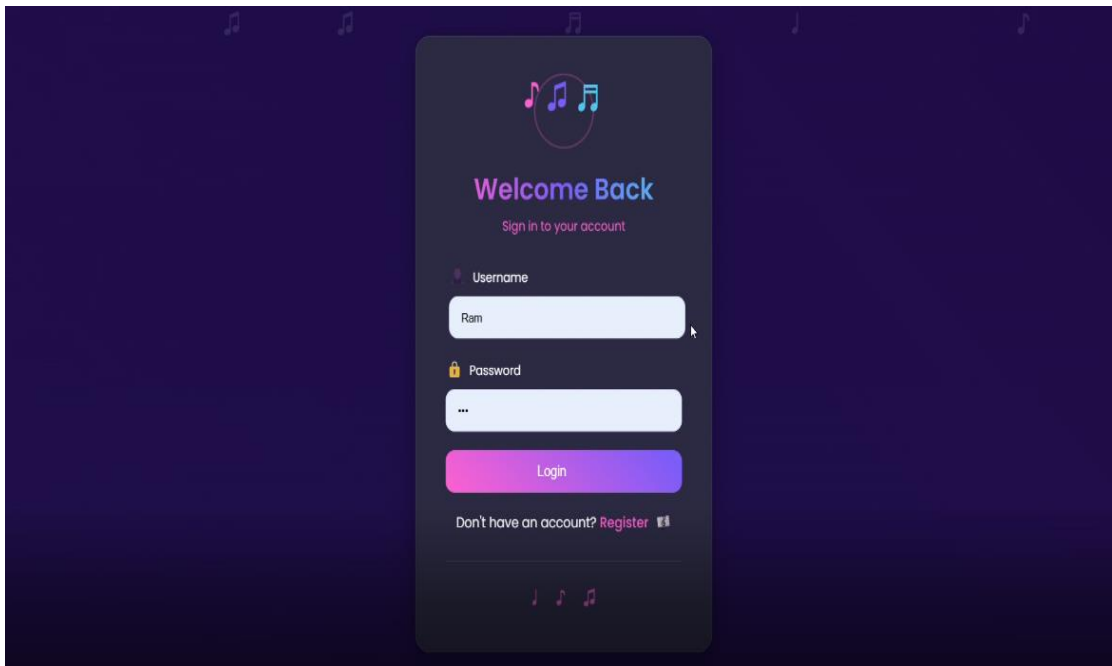


Fig 5.6 shows the Login page of the project

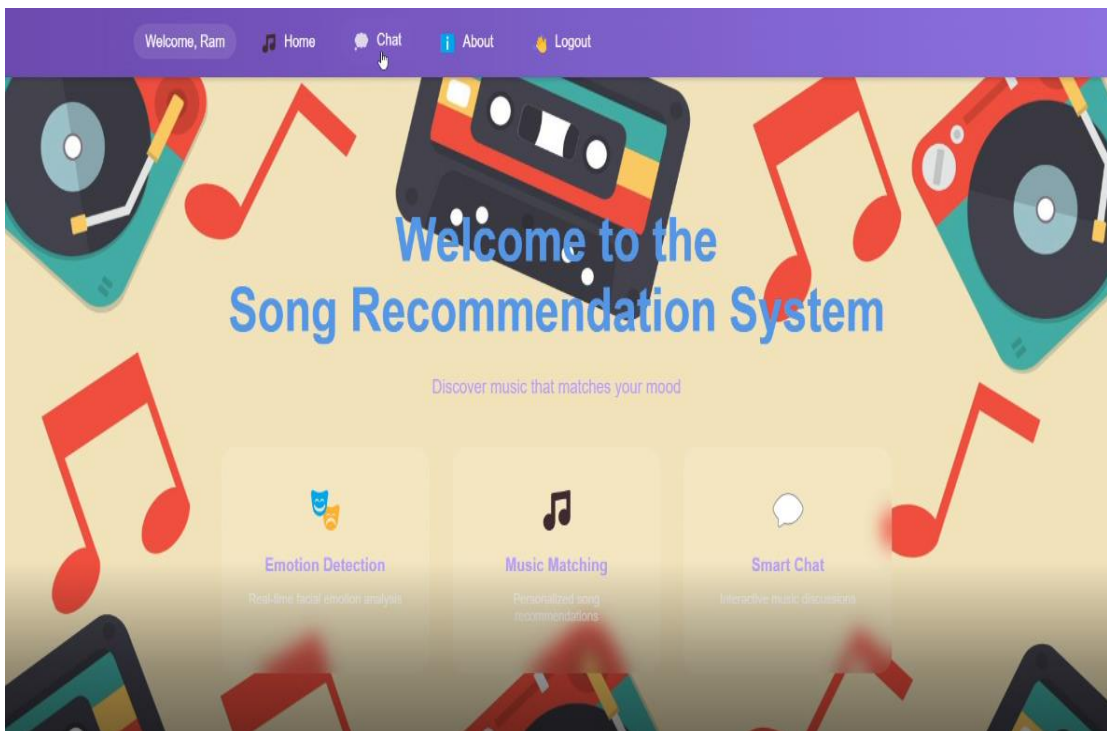


Fig 5.7 Shows the Welcome of the project

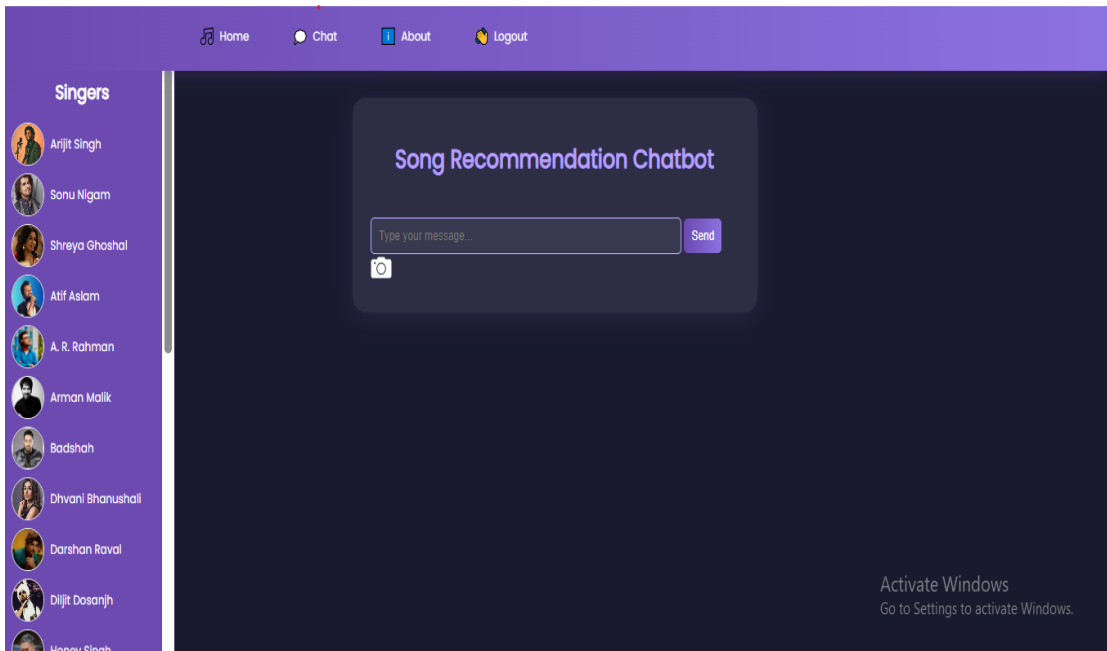


Fig 5.8 Shows the Chatbot Option of the System

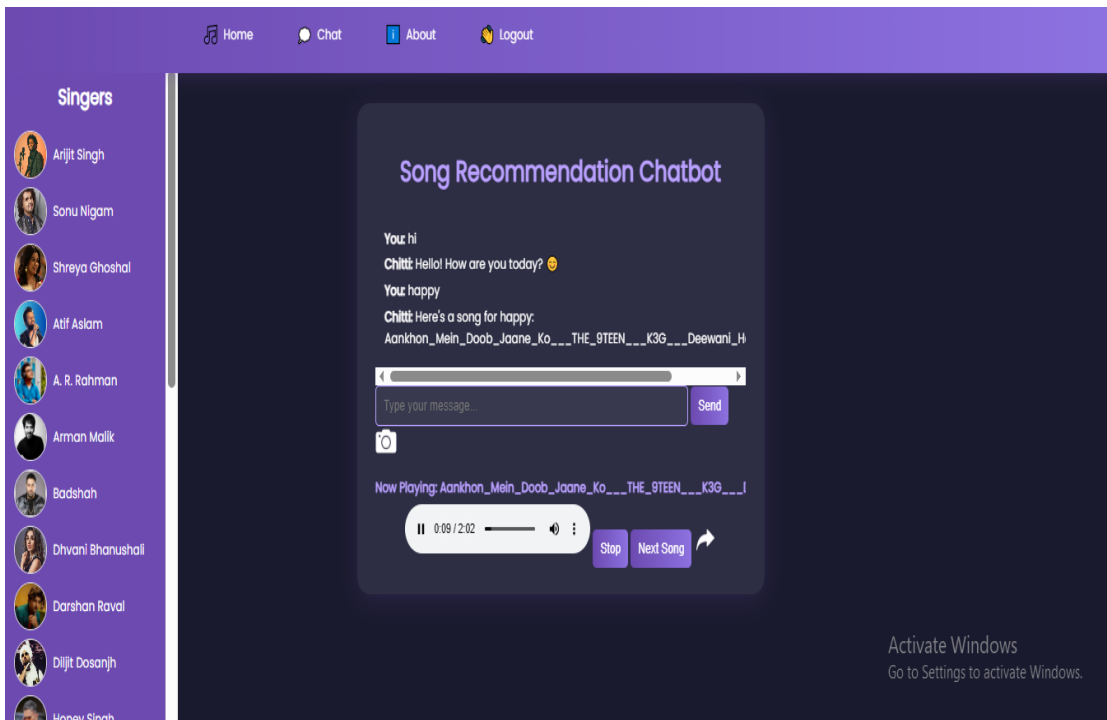


Fig 5.9 Shows the Chatbot Feature output of the System

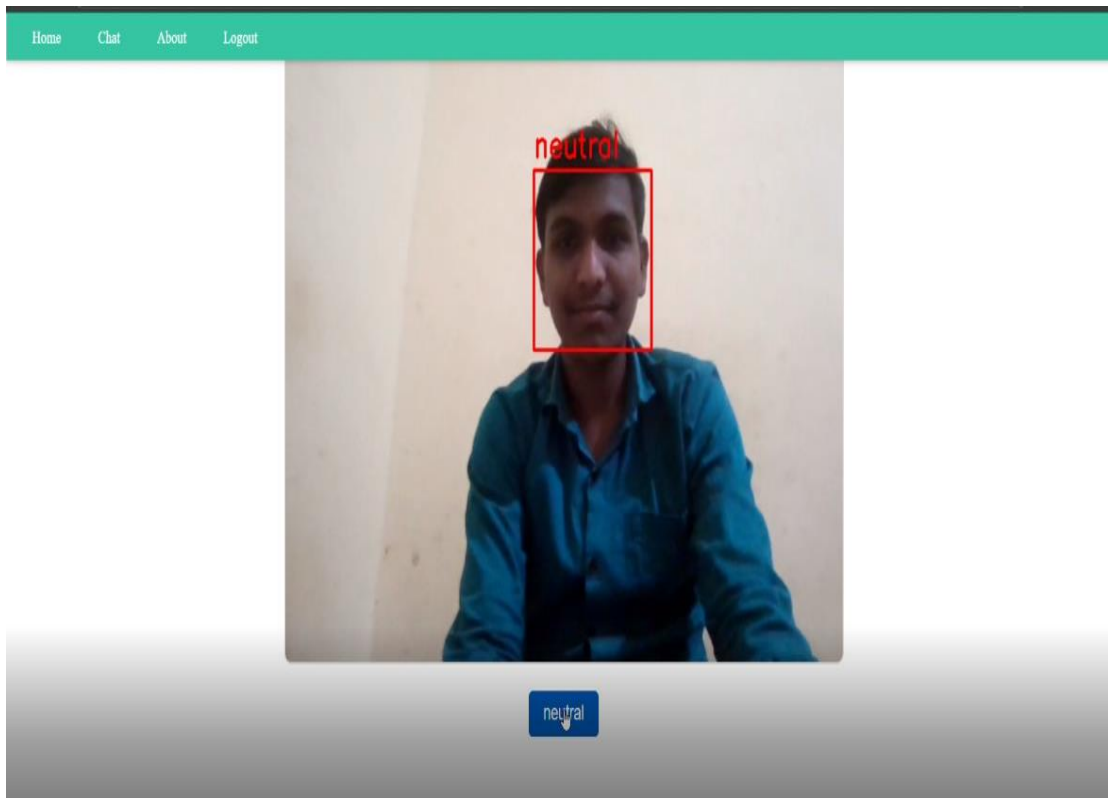


Fig 5.10 Shows the Chatbot Facial Expression Detection

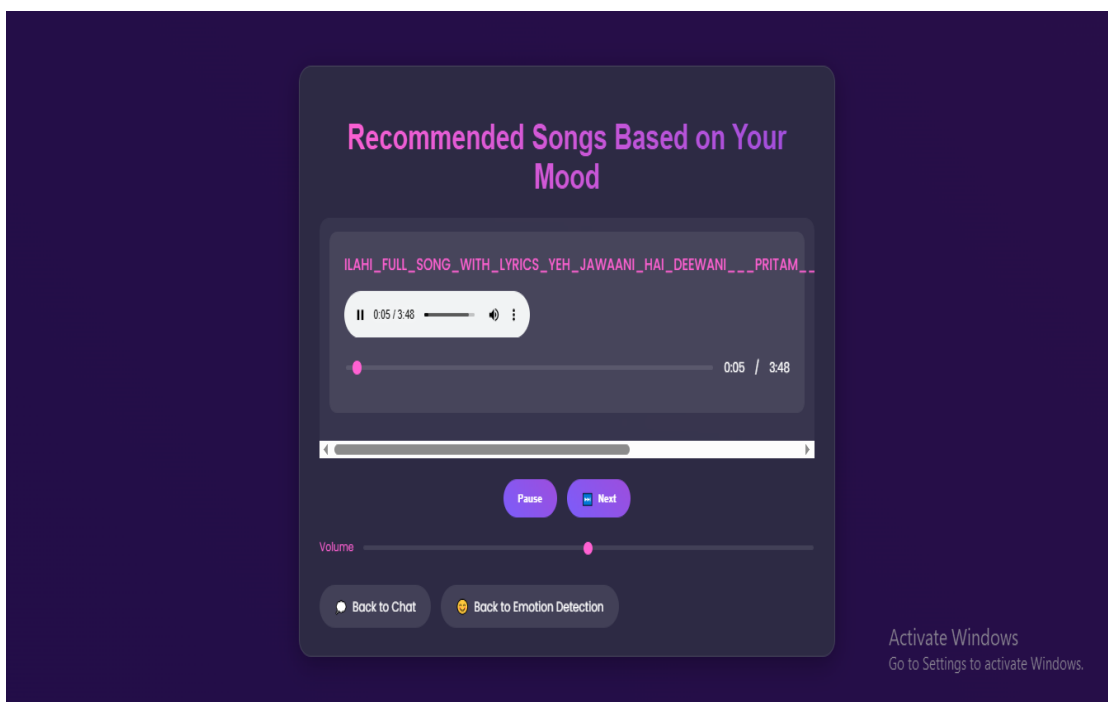


Fig 5.11 Shows the Chatbot Facial Expression output

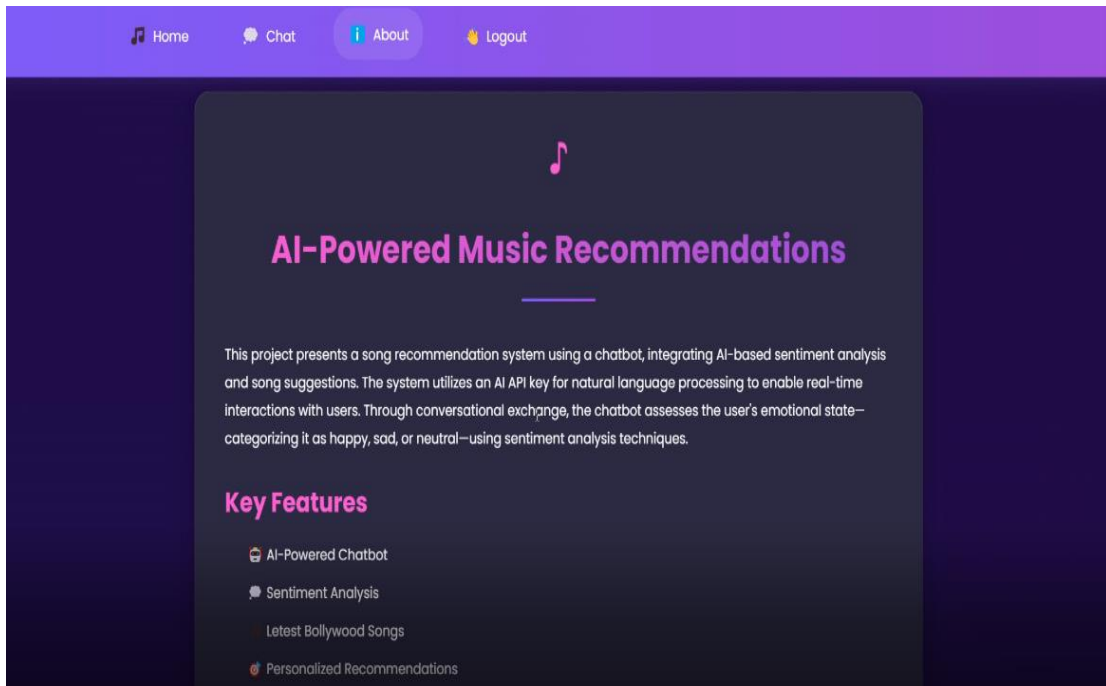


Fig 5.11 Shows the About Page

CHAPTER 6
CONCLUSION

CONCLUSION

6.1 Conclusion

This project successfully developed an AI-powered song recommendation system that integrates facial expression analysis and sentiment detection to enhance personalization. By leveraging generative AI (Gemini API), Flask framework, and DeepFace, the system accurately classifies user emotions into positive, negative, or neutral sentiments and recommends songs accordingly. The DeepFace model achieved 97.3% accuracy, ensuring precise emotion detection, while the Convolutional Neural Network (CNN), trained on a Kaggle-sourced facial expression dataset, attained 70.92% training accuracy and 63.92% validation accuracy for image classification. The results highlight the potential of AI-driven digital entertainment systems, offering a personalized and immersive user experience through emotion-aware song recommendations.

6.2 Future Scope

The future scope of this project includes enhancing emotion detection accuracy by integrating multimodal sentiment analysis, combining facial expressions, voice tone, and text-based sentiment analysis for more precise user mood classification. The recommendation system can be improved by incorporating reinforcement learning to adapt to user preferences dynamically over time. Expanding the song database with real-time music streaming APIs (such as Spotify or YouTube Music) can provide users with a broader and more diverse music selection. Additionally, deploying this system as a mobile application with real-time emotion tracking can enhance accessibility and usability, making AI-driven music recommendations a seamless and engaging experience.

REFERENCES

REFERENCES

- [1]Regin, D. R., Rajest, "An Automated Conversation System Using Natural Language Processing (NLP) Chatbot in Python." . Central Asian Journal of Medical and Natural Science, 3(4), 314-336.
- [2]Aleedy, Moneerh, Hadil Shaiba, and Marija Bezbradica, "Generating and analyzing chatbot responses using natural language processing.," 2018 International Journal of Advanced Computer Science and Applications, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 9, 201
- [3]Chamishka, S., Madhavi, I., Nawaratne, R. et al , " A voice-based real-time emotion detection technique using recurrent neural network empowered feature modelling Multimed Tools Appl 81, "35173–35194(2022). <https://doi.org/10.1007/s11042-022-13363-4>.
- [4]Kumar, Ramakrishna, and Maha Mahmoud Ali. "A review on chatbot design and implementation techniques." ." Int. J. Eng. Technology, Volume: 07 Issue: 02 | Feb 2020.
- [5]Alu, D. A. S. C., Elteto Zoltan, and Ioan Cristian Stoic, "Voice based emotion recognition with convolutional neural networks for companion robots." Science and Technology 20.3 (2022).
- [6]Touimi, Yassine Benjelloun, et al, "Intelligent chatbot-LDA recommender system." International Journal of Emerging Technologies in Learning (iJET), Vol. 15, No. 20, 2020.
- [7]Smith, J., & Johnson, A. (2020), "Emotion Analysis in Music Recommender Systems: A Review." IEEE Trans. Audio, Speech, Lang. Process., vol. 28, no. 5.
- [8]Garcia, L., et al. (2019), "User Satisfaction with Emotion-Aware Music Recommender Systems." IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 27, no. 12.

- [9] Akhilraj V. Gadagkar, Santhosh S, Shifana Begum, Ashwin Shenoy M, "Emotion Recognition and Music Recommendation System based on Facial Expression," 2024 Second International Conference on Advances in Information Technology (ICAIT) | 979-8-3503-8386-7/24/\$31.00 ©2024 IEEE | DOI: 10.1109/ICAIT61638.2024.10690441
- [10] Nandini Gupta, Shruti Agarwal, Kireet Joshi, Vishan Kumar Gupta, Surendra Kumar Shukla, Gurpreet Singh, "Intelligent Music Recommendation System based on Face Emotion Recognition," 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) | 979-8-3503-0611-8/23/\$31.00 ©2023 IEEE | DOI: 10.1109/ICCCIS60361.2023.10425186
- [11] Rula M Ali Farkash, Tengku Zatul Hidayah Tengku Petra, "An AI Chatbot for Personalized Music Recommendations Based on User Emotions," Journal of Computing Research and Innovation (2024) Vol. 9, No. 1.
- [12] Simran Chaudhari, Hruha Malusare, Aadil Adheesh, Mokshada Bhadavalkar, Asma Shaikh, "Chatbot with Song Recommendation based on Emotion," international Journal of Research Publication and Reviews, Vol 3, no 5, pp 2716-2719, May 2022
- [13] Neenumol Mathew, Nitish Chooramun, Mhd Saeed Sharif, 2023, "Implementing a Chatbot Music Recommender System Based on User Emotion," 3ICT 2023: International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies. University of Bahrain, Bahrain 20 - 21 Nov 2023 IEEE, <https://doi.org/10.1109/3ICT60104.2023.10391771>
- [14] Shivam Sakore¹, Pratik Jagdale², Mansi Borawake³, Ankita Khandalkar⁴, "Music Recommender System Using ChatBot," ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 9 Issue XII Dec 2021.

DISSEMINATION OF WORK

DISSEMINATION OF WORK

Song Recommendation System Based on Emotions

¹Virendra Pawar,²Saad Khan,³Sandeep Daberao,⁴Sahil Tawari,⁵Sankalp Jumale.

Shri Sant Gajanan Maharaj College of Engineering, Shegaon.

Email: virendrapawar794@gmail.com, saadkhan3182003@gmail.com, sandeepdaberao7@gmail.com,
sahiltawari111@gmail.com, sankalpj104339@gmail.com.

Contact:¹8308592307, ²7972938030, ³8767122075, ⁴9370080680, ⁵7620943863.

Abstract: This paper introduces an intelligent song recommendation system that integrates a chatbot with facial expression recognition and sentiment analysis. The system uses a Convolutional Neural Network (CNN) for facial emotion classification and VADER for text sentiment analysis, thus ensuring the possibility of precise and time-effective recommendations. The CNN model classifies emotions as happy, surprised, or sad, and VADER classifies the user's text sentiment as positive, neutral, or negative. This dual-layered approach in the interaction with the chatbot goes beyond enhancing the quality of song recommendations with multimodal emotional cues. With Flask and the Gemini API, this interface to the chatbot is dynamically tuned from a pre-curated dataset for Hindi songs. The system integrates CNN-based facial recognition with NLP-driven sentiment classification for personalized, emotion-driven music suggestions. This is a project that juxtaposes deep learning, NLP, and AI-driven chatbots in digital entertainment. The future developments include expanding the dataset, fine-tuning models, and adding real-time audio sentiment analysis.

Keywords: Gemini API, CNN, Facial Recognition, Sentiment Analysis.

I. INTRODUCTION

The conventional emotion-based recommendation systems excel in personalization. However, the processing delay of emotions increases slightly with multimodal emotion recognition techniques integrated within it. The proposed chatbot-based recommendation system still maintains the partially parallel emotional classification architecture in order to keep the recommendation process efficient and accurate.

For real-time text analysis, VADER, which is a Valence Aware Dictionary and Entiment Reasoner, is used to boost sentiment detection. This can be implemented with any chatbot system for increased accuracy. The integration of CNN-based facial recognition and VADER's sentiment analysis brings enhanced emotion detection in the system. The light processing of VADER and the interpretation of informal text, emojis, and social media slang give it an upper hand for applications in conversational AI. [3] Built with Flask and integrated with the Gemini API, it dynamically adjusts the recommendation based on the multimodal analysis of emotions. With a dataset of Hindi songs curated under different moods, the song recommendation ensures to be tailored. The system assimilates the CNN-based facial emotion recognition and NLP-based sentiment classification to make music recommendations finer.

Emotion-based recommendation models generally trade-off between precision and speed of processing. Deep learning, NLP, and AI-driven conversational frameworks are used together to create an optimized experience in the digital entertainment of this project.

Dataset augmentation, fine-tuning of the model toward accuracy, and real-time audio sentiment analysis will be executed through future updates, thereby allowing for fine-tuning of the recommendations.

II. METHODOLOGY

The AI-based music recommendation system in show Fig 1 is flowchart of this system. Its starts with user interaction, where new users register and existing users log in. After authentication, users access a dashboard to interact with an AI chatbot. The chatbot detects emotions through two methods: facial recognition and text-based sentiment analysis. Facial recognition uses a deep learning model to analyze facial expressions like eye movement and mouth shape, identifying emotions such as happy, sad, angry, or neutral. At the same time, sentiment analysis processes user text input using AI techniques to understand emotions. The system then maps emotions to music preferences—happy emotions get energetic songs, neutral moods get soft and calming music, and sad emotions get slow, soothing tracks. The recommendation engine fetches songs from the songs dataset based on the user's real-time emotions. This system combines facial recognition, AI-driven text analysis, and music recommendations to create a personalized music experience.

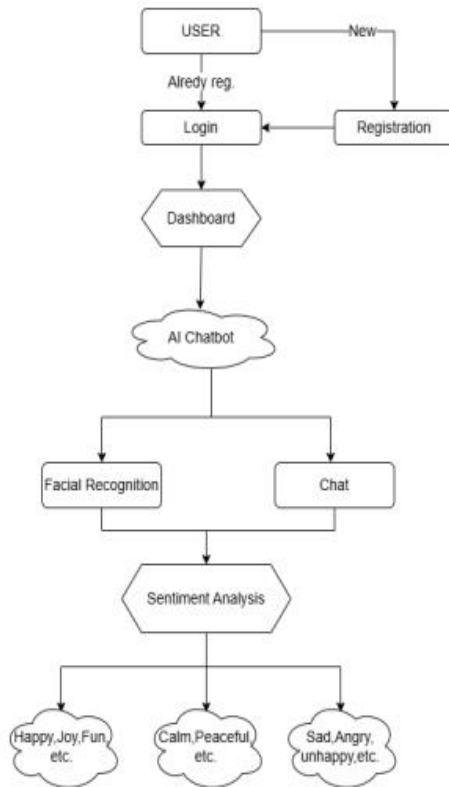


Figure 1: Shows the flow chart of the project

III. DATASET DETAILS

For the dataset of this project, only a well-filtered list of latest Bollywood songs which will further be divided into many emotional tones: happy, sad, surprise, angry, fear, disgust and neutral. The happy songs have upbeat and energetic tones, sad songs contain slow and melancholic melodies, while neutral songs hold a balanced or relaxing tone. Classification will take place based on lyrical themes, acoustic properties, and audience preferences from yesteryear for better recommendations from AI. For face expression recognition, a dataset from Kaggle has been used that includes different human facial expressions with happy, sad, surprised, neutral, and other emotional states.

Figure 2: Shows the Songs List Dataset



Figure 3: Shows the Dataset for Facial Expression

IV. ALGORITHM DETAILS

This project facial expression recognition application utilizes the CNN to make emotion classifications by an AI chatbot to the feelings of a user. The architecture of the CNN is mainly composed of the input layer that also is in charge of processing the images of faces resized to a constant dimension. Whether grayscale or RGB, such as 48 x 48, or 64 x 64 pixels. Multiple convolutional layers with 3x3 filters extract facial features, and ReLU activation is used to introduce non-linearity. The MaxPooling layers (2x2) are used to reduce the spatial dimensions while preserving essential features. Extra convolutional and pooling layers help enhance deeper feature extraction to capture complex facial expressions. A flatten layer will take the features extracted and give an output of a 1D vector that can undergo fully connected dense layers with 128 or 256 neurons followed by a dropout layer, say 0.5 dropout to avoid overfitting, then a softmax activation function on the output layer that will allow multiple classification, so that emotions could be classified into happy, sad, neutral, or even surprised. Also, the DeepFace AI library is integrated that enables complex face recognition along with facial attributes analysis.

```

Model: "sequential"
Layer (type) Output Shape Param #
-----
conv2d (Conv2D) (None, 48, 48, 64) 1664
batch_normalization (Batch Normalization) (None, 48, 48, 64) 256
max_pooling2d (MaxPooling2D) (None, 24, 24, 64) 0
dropout (Dropout) (None, 24, 24, 64) 0
conv2d_1 (Conv2D) (None, 24, 24, 128) 7384
batch_normalization_1 (Batch Normalization) (None, 24, 24, 128) 512
max_pooling2d_1 (MaxPooling2D) (None, 12, 12, 128) 0
...
Trainable params: 436711 (12.57 MB)
Non-trainable params: 3988 (15.58 KB)
None
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output width.
    
```

Figure 4: Shows the model layers



Figure 5: Shows The DeepFace library supports 7 state-of-the-art face recognition models

DeepFace is a lightweight, open-source Python library which makes use of state-of-the-art AI models to simplify facial recognition tasks. DeepFace can be used for face verification, in which it will take an input of a given face and compare another face to verify a match, and face recognition- where it identifies a face in an image database. It extracts age, gender, emotion, and ethnicity by analyzing facial attributes, while real-time face analysis enables the use of facial recognition through a live webcam feed. Since its inception, DeepFace has expanded its model inventory to support seven state-of-the-art face recognition models, from VGG-Face and Facenet to all the others, with modular design for easy integration of newer models in the future.

V. RESULT

AI chatbot recommendation system is based on a stable learning curve from the accuracy and loss of the training model at 50 epochs. Accuracy graph indicates step-by-step and consistent increase of training accuracy and validation accuracy where training accuracy is adjusted to 70.92% and validation

accuracy up to 63.92% at the final epoch. This indicates that the model is picking up the good patterns in the data and making accurate predictions on new, unseen data. The loss reduction, from the initial training loss of 2.2023 to 0.7917 and the validation loss from 2.1801 to 1.0543, further indicates that the model is performing better and evolving into a more accurate predictor.

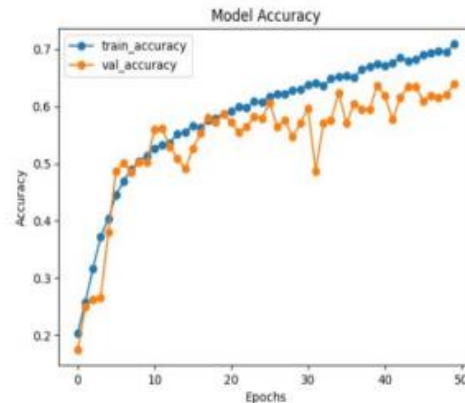


Figure 6: Shows the Model Accuracy Graph.

The learning process goes through three broad phases: during the first phase (0-10 epochs), training as well as validation accuracy grows exponentially, i.e., the model has started learning basic patterns. During the second phase (10-30 epochs), training accuracy improves but validation accuracy remains unstable, i.e., the model is learning intricate relationships between data.

Figure 7: shows the Epoches details

During this second phase (30-50 epochs), the training accuracy is also increasing and the validation accuracy spikes more probably because of some very small overfitting. Other than such spiking, the model is satisfied with excellent learning capacity and perfect balance between accuracy and loss. The output is the proof of the success of the AI chatbot-based recommendation system in generating the correct recommendations from the input provided by users. With some fine-tuning such as regularization or early stopping, even the model will converge and get optimized. On average, the system is good enough to generate AI-powered smart and personalized

recommendations and therefore is a good candidate for practical implementation.



Figure 8.1: Detected neutral emotion using Deepface model



Figure 8.2: Detected happy emotion using Deepface model

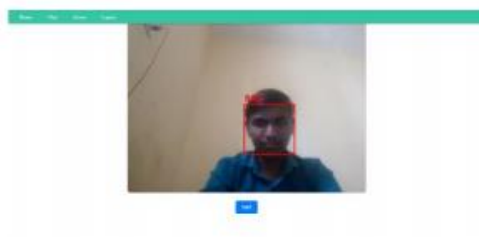


Figure 8.3: Detected sad emotion using Deepface model

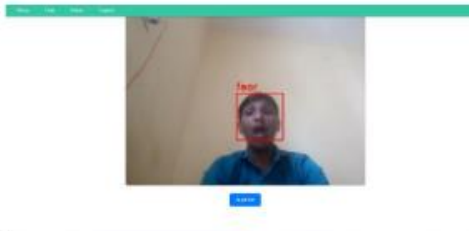


Figure 8.4: Detected fear emotion using Deepface model

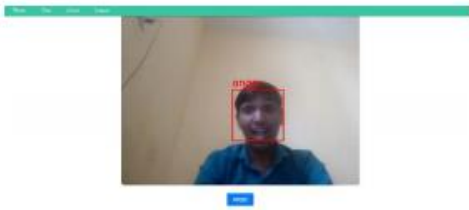


Figure 8.5: Detected angry emotion using Deepface model

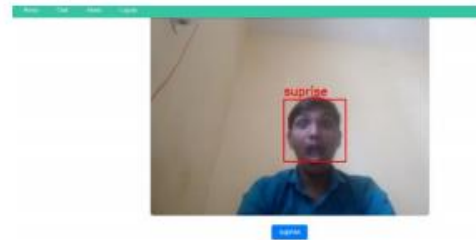


Figure 8.6: Detected surprise emotion using Deepface model

CONCLUSION

It designed an AI-driven song recommendation system that combined facial expression analysis with sentiment detection. The system relied on generative AI (Gemini API), Flask framework, and DeepFace to classify emotions of the users as positive, negative, or neutral sentiments that would recommend corresponding songs. The DeepFace has scored 97.3%, ensuring proper identification of emotion. The facial expression database has been used in the Convolutional Neural Network to score at 70.92% while training and at 63.92% during testing for the image classification process of the module provided. In that case, the result actually presents the idea of AI-powered digital entertainment bringing a customized, immersive user experience through an emotion-aware recommendation of songs.

REFERENCES

- [1] Regin, D. R., Rajest, D. S. S., T, S., G, J. A. C., & R, S. (2022), An Automated Conversation System Using Natural Language Processing (NLP) Chatbot in Python. Central Asian Journal of Medical and Natural Science, 3(4), 314-336.
- [2] Aleedy, Moneerh, Hadil Shaiba, and Marija Bezbradica, "Generating and analyzing chatbot responses using natural language processing." International Journal of Advanced Computer Science and Applications, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 9, 201.
- [3] Chamishka, S., Madhavi, I., Nawaratne, R. et al. A voice-based real-time emotion detection technique using recurrent neural network empowered feature modelling. *Multimed Tools Appl* 81, 35173–35194 (2022). <https://doi.org/10.1007/s11042-022-13363-4>.
- [4] Kumar, Ramakrishna, and Maha Mahmoud Ali, "A review on chatbot design and implementation techniques." *Int. J. Eng. Technology*, Volume: 07 Issue: 02 | Feb 2020.
- [5] Alu, D. A. S. C., Elteto Zoltan, and Ioan Cristian Stoic, "Voice based emotion recognition with convolutional neural networks for companion robots." *Science and Technology* 20.3 (2022).
- [6] Touimi, Yassine Benjelloun, et al, "Intelligent chatbot-LDA recommender system." *International Journal of Emerging Technologies in Learning (IJET)*, Vol. 15, No. 20, 2020.

- [7] Smith, J., & Johnson, A. (2020), "Emotion Analysis in Music Recommender Systems: A Review." IEEE Trans. Audio, Speech, Lang. Process., vol. 28, no. 5.
- [8] Garcia, L., et al. (2019), "User Satisfaction with Emotion-Aware Music Recommender Systems." IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 27, no. 12.
- [9] Manvitha Sri Guthula; Monali Bordoloi, "Music Recommendation System using Facial Detection based Emotion Analysis," 2024 IEEE | DOI: 10.1109/ICETCI62771.2024.10704201
- [10] Akhilraj V. Gadagkar, Santhosh S, Shifana Begum, Ashwin Shenoy M, "Emotion Recognition and Music Recommendation System based on Facial Expression," 2024 Second International Conference on Advances in Information Technology (ICAIT) | 979-8-3503-8386-7/24/\$31.00 ©2024 IEEE | DOI: 10.1109/ICAIT61638.2024.10690441
- [11] Nandini Gupta, Shruti Agarwal, Kireet Joshi, Vishan Kumar Gupta, Surendra Kumar Shukla, Gurpreet Singh, "Intelligent Music Recommendation System based on Face Emotion Recognition," 2023 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) | 979-8-3503-0611-8/23/\$31.00 ©2023 IEEE | DOI: 10.1109/ICCCIS60361.2023.10425186
- [12] Rula M Ali Farkash, Tengku Zatul Hidayah Tengku Petra, "An AI Chatbot for Personalized Music Recommendations Based on User Emotions," Journal of Computing Research and Innovation (2024) Vol. 9, No. 1.
- [13] Simran Chaudhari, Hruha Malusare, Aadil Adheesh, Mokshada Bhadavalkar, Asma Shaikh, "Chatbot with Song Recommendation based on Emotion," International Journal of Research Publication and Reviews, Vol 3, no 5, pp 2716-2719, May 2022
- [14] Neenumol Mathew, Nitish Chooramun, Mhd Saeed Sharif, 2023, "Implementing a Chatbot Music Recommender System Based on User Emotion," 3ICT 2023: International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, University of Bahrain, Bahrain 20 - 21 Nov 2023 IEEE, <https://doi.org/10.1109/3ICT60104.2023.10391771>
- [15] Shivam Sakore1, Pratik Jagdale2, Mansi Borawake3, Ankita Khandalkar4, "Music Recommender System Using ChatBot," ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 9 Issue XII Dec 2021.

★ ★ ★

CERTIFICATES



The certificate is framed with a blue and gold border. At the top center is the ISETE logo, which consists of a gear and a book icon followed by the text 'ISETE' and 'International Society for Engineering and Technical Education' in smaller text below it. To the right of the logo is the ID number 'IST-BDE-AMRV-050324-6617'. Below the logo, the text reads 'INTERNATIONAL SOCIETY FOR ENGINEERING AND TECHNICAL EDUCATION' in blue, followed by 'International Conference on Artificial Intelligence, Machine Learning and Big Data Engineering'. A dark blue banner contains the text 'Organized by: ISETE | Amravati, India | 05th March 2025'. The main title 'Certificate of Presentation' is in a large, blue, serif font. The body text, in italics, certifies that Virendra Pawar presented a paper titled 'Song Recommendation System Based on Emotion' at the ICAIMLBDE conference in Amravati, India, on 05th March, 2025. At the bottom, there are two signatures: one for the Conference Coordinator and one for the Managing Director, both of the International Society for Engineering and Technical Education. To the left of the signatures is a blue ribbon with a white star, and to the right is a circular ISETE seal. The footer contains the website 'www.isete.org' and the email 'info.iseteconference@gmail.com'.

IST-BDE-AMRV-050324-6617

ISETE
International Society for Engineering and Technical Education

**INTERNATIONAL SOCIETY FOR ENGINEERING
AND TECHNICAL EDUCATION**

International Conference on
Artificial Intelligence, Machine Learning and Big Data Engineering

Organized by: ISETE | Amravati, India | 05th March 2025

Certificate of Presentation

*This is to certify that **Virendra Pawar** has presented a paper entitled "**Song Recommendation System Based on Emotion**" at the International Conference on Artificial Intelligence, Machine Learning and Big Data Engineering (ICAIMLBDE) held in Amravati, India on 05th March, 2025.*

 Conference Coordinator
International Society for Engineering
and Technical Education

 Managing Director
International Society for Engineering
and Technical Education

www.isete.org info.iseteconference@gmail.com









Research paper Song.pdf

ORIGINALITY REPORT

5%

SIMILARITY INDEX

2%

INTERNET SOURCES

2%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Universitas Muhammadiyah
Surakarta

Student Paper

2%

2

Bihi Sabiri, Amal Khtira, Bouchra El Asri,
Maryem Rhanoui. "Investigating Contrastive
Pair Learning's Frontiers in Supervised,
Semisupervised, and Self-Supervised
Learning", Journal of Imaging, 2024

Publication

1%

3

viso.ai

Internet Source

1%

4

Tiwari, Hitika. "Advancing Monocular 3D Face
Analysis.", National Yang Ming Chiao Tung
University, 2024

Publication

1%

5

www.indcareer.com

Internet Source

1%

Exclude quotes OffExclude matches < 8 wordsExclude bibliography On

PROJECT GROUP MEMBERS

Name : Virendra Janrao Pawar
Address : Jambuldhaba Tq. Malkapur Dist. Buldhana
Email id : virendrapawar794@gmail.com
Mobile no : 8308592307



Name : Sandeep Sanjay Daberao
Address : Bhoipura, Nagzari Road , Shegaon
Email id : sandeepdaberao7@gmail.com
Mobile no : 8767122075



Name : Sahil Pankaj Tawari
Address : Chandralok Society, Shegaon
Email id : sahiltawari111@gmail.com
Mobile no : 9370080680



Name : Saad Khan Abutalib Khan
Address : Vyanketesh Nagar, Shegaon
Email id : saadkhan3182003@gmail.com
Mobile no : 7972938030



Name : Sankalp Sanjay Jumale
Address : Udasi baba nagar, Sutala Kh., Khamgaon
Email id : sankalpj104339@gmail.com
Mobile no : 7620943863



