IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EMOTUNE

Emotion and Gender Aware Music Generation Chatbot

[1]PROF. ARCHANA PRIYADARSHINI RAO, [2]NITHASHA, [3]PRATHEEKSHA R, [4]PRATHIK S, [5]RACHANA RAO

[1]Faculty, [2]Student, [3]Student, [4]Student, [5]Student

[1]Department of Information Science and Engineering,

[2]A J Institute of Engineering and Technology, Mangalore, Karnataka, India

[3]A J Institute of Engineering and Technology, Mangalore, Karnataka, India

[4]A J Institute of Engineering and Technology, Mangalore, Karnataka, India

[5]A J Institute of Engineering and Technology, Mangalore, Karnataka, India

Abstract: : In an era where technology is continually shaping our interactions with the world, "EmoTune" emerges as an innovation at the crossroads of music and artificial intelligence. This project represents a step in the evolution of human-computer interaction, introducing an extraordinary chatbot system capable of crafting music that transcends language barriers and connects deeply with human emotions. With a user-friendly interface designed to accept real-time text, video, audio, and image inputs, it is poised to redefine the way we experience and create music. At its core, it takes inspiration from the universal language of music and blends it with the power of advanced emotion and gender recognition technology. It possesses the remarkable ability to generate English lyrics for songs in any language, opening the doors to a world of cross-cultural musical storytelling. As users engage with the chatbot, their emotions are detected and analyzed in real time, ensuring that the music created is not just linguistically accurate but also emotionally resonant. This empathetic connection between technology and emotion transforms It into a personalized musical companion, capable of crafting melodies that mirror the user's feelings and preferences. Furthermore, it pays homage to the rich heritage of music by delving into the timeless emotions of classic songs. It extends its creative reach to evoke the nostalgia associated with old songs, acknowledging the unique emotional journeys of older generations. In doing so, it breathes new life into these cherished melodies, offering users an opportunity to revisit and rediscover the sentiments of the past. In the ever-evolving landscape of artificial intelligence, ethical considerations remain paramount. It seeks to enrich the human-computer interface by offering a deeply personal and culturally diverse musical experience, where emotions and melodies converge to inspire and uplift. "EmoTune: Emotion and Gender-Aware Music Generation Chatbot" is poised to become a transformative force, ushering in a new era of harmonious human- technology interaction.

Index Terms **- Music generation, Chatbot, Emotion, Gender, Audio, Video, Text**

I. INTRODUCTION

"EmoTune: Emotion and Gender-Aware Music Generation Chatbot" heralds a new era at the crossroads of artificial intelligence and musical expression, pioneering an innovative approach to personalized soundscapes. In a digital landscape saturated with music platforms, it distinguishes itself by dynamically responding to user emotions and gender preferences in real-time. Whether it's the infusion of sentiment from textual inputs, the interpretation of visual cues through images, or the subtle nuances captured in audio and video, it transforms the user's emotional state into a unique musical journey. The chatbot's adaptability goes beyond the conventional; it incorporates a thoughtful feature that recognizes and caters to the musical tastes of the older generation. Upon identifying an aged user, it gracefully transitions into generating classic songs, paying homage to timeless melodies that resonate with the rich cultural heritage of bygone eras. This distinctive touch not only adds a nostalgic dimension to the musical repertoire but

also acknowledges the unique emotional journeys of those who have witnessed the evolution of musical styles over the years.

At its core, it is a user-friendly interface designed to transcend generational boundaries and provide a transformative musical companion. Its versatile architecture accommodates diverse forms of user interaction, offering a seamless blend of emotional resonance and cultural inclusivity. Whether one is a music enthusiast or a seasoned developer, it stands as an accessible and innovative force in the realm of AI-driven music chatbots.

As we navigate an increasingly interconnected world, it emerges as a beacon of creativity and empathy, reshaping the contours of musical engagement. This project reflects a commitment to harnessing technology not just for utility but for fostering deeper connections, celebrating emotions, and recognizing the varied musical tastes that span generations. it invites users on a dynamic and personalized journey through the universal language of music, where every note resonates with the intricacies of human emotion and cultural diversity.

II. PROBLEM STATEMENT

The contemporary digital music landscape lacks a personalized and emotionally resonant experience for users. Existing platforms often fall short in understanding and responding to the intricate nuances of individual emotions and gender preferences. The existing system is restricted to a single mode. This unimodal approach overlooks the diverse ways users express their emotions and preferences. This rigid structure fails to capture the richness of human expression, limiting the depth and personalization of the music recommendation and generation processes. The absence of a dynamic and adaptive music generation system limits users to static playlists, overlooking the potential for a more immersive and tailored musical journey. Additionally, the emotional and cultural connections embedded in classic songs often go unexplored, neglecting the rich heritage and unique emotional landscapes of older generations.

III. OBJECTIVES

- Enable users to interact with EmoTune in real-time through multimodal inputs for a responsive experience.
- Implement algorithms to tailor music compositions based on user gender preferences, creating a personalized and inclusive musical experience.
- Implement a feature that identifies aged users and generates classic songs to provide a nostalgic musical journey.
- Enable to generate English lyrics for songs in any language.
- Incorporate ethical considerations in the design and implementation of EmoTune, ensuring user privacy.
- Develop EmoTune to handle diverse input types and ensure compatibility with popular development environments for widespread usability.

IV. LITERATURE SURVEY

This chapter summarizes the work related to Emotion and Gender Aware Music Generation Chatbot, that are carried out by different researchers.

1. **“Chatbot Song Recommendation System”,** 2023 by Anusha, Dr. Srinivasan V:
   It recommends appropriate music based on the user's emotional state using a suitable API that is readily available. This paper focuses on a chatbot that uses artificial to analyse the user's tone in text form, as users increasingly interact with systems through text and voice assistants. Here, chatbot saves the user's input after each interaction and uses a tone analyzer API to process the user's text response for subsequent interactions. They used Last.fm API to retrieve music information. As a result, the dataset contains instances of exchanges and facial expressions that

communicate emotions in all three directions: Positive (happy, excitement, or joy), Negative (sadness, wrath, or frustration), and Neutral (neutrality, indifference, or objectivity). While current research primarily emphasizes response improvement, there is a pressing need to explore linguistic aspects like emotional and sentiment analysis. Moreover, the proposed work on human emotion recognition can be extended to effectively recognize mixed emotions, enabling a deeper understanding of individuals' emotional states.

2. **"Chatbot with Song Recommendation based on Emotion",** 2022 by Simran Chaudhari, Hrucha Malusare, Aadil Adheesh, Mokshada Bhadavalkar, Asma Shaikh**:**

They implemented a chatbot that interacts with the user, analyses the emotions of chats, and recommends a song playlist based on the user's emotions. Their objective was to identify the emotion perceived by the user, and once the emotion is identified, a list of songs is suggested based on the emotion. They had built a simple retrieval based chatbot which uses predefined input patterns and responses. They had used a special recurrent neural network named Long Short-Term Memory to classify the category of a user's message and get an appropriate response from the chatbot. Support Vector Machine, Linear Support Vector Machine, Random Forest, and Decision Tree Classifiers are used to detect the emotions expressed by the user in chats. The advantage of preserving the model is that you don't have to fit the model every time you execute the code, which would take a long time. The load function can be used to load the model. One drawback of this approach is the potential lack of adaptability. Predefined input patterns and responses mean that the chatbot might struggle to handle new or unexpected user messages that fall outside its predetermined patterns.

3. **"Mood Based Music Playlist Generator Using Convolutional Neural NetworkProf"**, 2022 by Jaychand Upadhyay, Sharan Shetty, Vaibhav Murari, Jarvis Trinidade:

This paper aimed to solve the issue of the manual finding of songs to suit one's mood along with creating a high accuracy CNN model for Facial emotion recognition. Through the webcam, the emotional state can be deduced from facial expressions. To create a neural network model, the CNN classifier was used. This model is trained and tested using OpenCV to detect mood from facial expressions. A playlist will be generated through the system according to the mood predicted bythe model. The model used for predicting the mood obtained an accuracy of 97.42% with 0.09 loss on the training data from the FER2013 Dataset. Advantage of using this is that the old or the current system requires the user to manually go and search for a playlist based on his mood, but here the proposed system makes this task very simple by just capturing an image of the user and detecting the user's mood and generating a playlist according to the mood detected to enhance the user's mood. One disadvantage is that it focuses only on emotions from facial expressions, but it might not always get the music right. Music tastes can be very personal, even if you feel a certain way, the songs suggested might not match what you really like.

4. **"Alexis: A Voice based Chatbot using Natural Language Processing"**, 2022 by Meet Popat, Aayush Doshi, Yashraj Rai:

It uses NLP for creating a conversational chatbot and uses natural language processing (NLP) in the form of Python code for integration using the Flask framework to enhance self-learning and 24/7 resource availability. The Natural Language Toolkit (NLTK) library was used to understand user input. It Works in 2 modes: Text (speech-to-text) and Speech(text-to-speech). AICTE.txt and SFIT.txt are the datasets used. The overall accuracy of the web API's response will be measured using a developer made testing set. Benefit of using chatbot is that it's efficient as it automates and simplifies user interaction and make it faster to provide information by recognizing the text input sent by the user and it is user-friendly. Chatbot can also improve their response by learning from user interaction and gradually becoming more accurate. Despite using natural language processing, the Chatbot may struggle to understand complex queries, which lead to incorrect responses.

5. **"Emotion Recognition Using Chatbot System"**, 2022 by Shraddha Pophale, Hetal Gandhi and Anil Kumar Gupta:

This system is built to recognize the emotions from the text conversations done by students. The sentiment viewpoints are recognized as the positive and negative descriptive words and different weight is assigned to these descriptors dependent on the criticality vectors. After assigning these weights, the following work is to process the single message. This procedure includes the sentence adjustment and recognizable proof of area angles and the sentiments. In light of the aspect and adjective relationship, the weight of a specific sentence is recognized. At the last stage, order

approach will be applied to identify the characteristics specific sentiment. LSTM Encoders and Decoders for Chatbot System and the output text processed for Emotion Class Prediction. It figures out how people are feeling from what they write by using chatbot, so it can respond in a way that matches those feelings. One drawback is that there is no privacy, where user feels like their personal information is being looked at without their permission.

6. **"A Machine Learning Based Chatbot Song Recommender System"**, 2021 by Amaan Shaikh, Bhushan Patil, Tejas Sonawane:

The proposed research work develops a personalized system, where the user's current emotion is analyzed with the help of the chatbot. The chatbot identifies the user's sentiment by asking some general questions. Based on the input provided by the user, current emotion or mood is analyzed by the chatbot and it will generate the playlist. It uses IBM Tone Analyzer API, Last.fm API, CakeChat server. Chat-bots have 24/7 Availability, Chat-bots cause an Increase in Sales, Chat-bots provide Long-term Financial Savings. But it has several drawbacks too, i.e, Chat-bots can only handle basic questions, Chat- bots are difficult to create, this requires complex programming and is not easy for companies. This becomes especially difficult if companies have to create chat-bots from scratch and that is why many online platforms help companies to build and manage chat-bots easily. Chat-bots require constant maintenance.

7. **"Music Recommender System Using ChatBot"**, 2021 by Shivam Sakore, Pratik Jagdale, Mansi Borawake, Ankita Khandalkar:

They have implemented a chatbot that recommends music based on the user's text tone. By analysing the tone of the text expressed by the user, it could identify the mood. Once the mood is identified, the application would play songs in the form of a web page based on the user's choice as well as his current mood. They used IBM Analyser to check the text tone of the user and to predict the mood based on the text of the user, and Last.FM API to recommend songs based on the mood of the user. They used number of pre-existing DSP algorithms are used in this analysis such as the fast Fourier transform (FFT), which displays the frequencies present in a time-domain signal. The tone analyzer is really good at figuring out how someone feels by looking at the words they write. It knows if someone is happy, sad, angry, or can recognize other emotions too. Despite using tools like IBM Analyzer, there could still be difficulties in predicting or understanding a user's mood solely from their text, which will lead to incorrect music selections.

8. **"A Novel Music Emotion Recognition Model Using Neural Network Technology"**, 28 Sep 2021 by Jing Yang:

An improved back propagation (BP) algorithm neural network is used to analyze music data. This paper introduces artificial bee colony (ABC) algorithm to improve the structure of BP neural network. The output value of the ABC algorithm is used as the weight and threshold of the B neural network. The ABC algorithm is responsible for adjusting the weights and thresholds, and feeds back the optimal weights and thresholds to the BP neural network system. Through experiments on public music data sets, the experimental results show that, compared with other comparative models, the MER method used in this paper has better recognition effect and faster recognitions. In order to improve the recognition rate of music emotion, this research uses an improved BP network to recognize music data. This research first classifies the acoustic features of music in a combined form for emotion classification, and analyzes the most suitable feature data for emotion recognition. Secondly, a music sentiment classifier was constructed using the BP network optimized by the ABC algorithm and compared with the experimental results of other classifiers. Based on the experimental results, it can be seen that the network used has a better recognition effect.

9. **"Music recommendation system based on facial emotion recognition"**, March 2020 by Deny John Samuel, B. Perumal, Muthukumaran Elangovan:

The Model recognizes the mood of the user from facial emotions and recommend music. The proposed system is both time and cost-efficient. Due to the unbalanced nature of each element set, effective approaches to incorporate different highlights and functionalities should be investigated. To improve the exactness of the arrangement framework the informational collection used to construct the grouping model could be expanded further. We also plan to develop some special features where quotation of some great personalities will be recommended based on the user emotions if the user is detected as sad a song was suggested according to that song a quote

will be displayed so that the user can be motivated. Each convolutional layer applies one new channel to the tangled image of the past layer that can eliminate one more part. Accordingly, as more channels are stacked, the more features the CNN can remove from an image. The three components that go into the convolution activity are Input image, Feature detector and Feature map.

10. **"Conversational and Image Recognition Chatbot"**, 2020 by Shengyang Su:
The model for the image reorganization chatbot is based on a encoder-decoder framework, composed of two parts 1. an encoder that converts the input into a vector space, 2. a decoder that converts the embedded vector into an output. Here, Late Fusion (LF) Encoder, Generative (LSTM) decoder and Discriminative (softmax) decoder are used. Localization of object or persons is done using Mask-RCNN model which not only localizes the object, but also provides a mask for localized object. The main purpose of this project is to utilize natural language processing and computer vision models for efficient identification and answer following question to any images. COCO dataset is being used to train the data and images are fused together to get more informative output for detection of a doubtful presence, but making these systems work together needs lots of expertise and can be hard to set up.

11. **"A Novel Music Emotion Recognition Model for Scratch generated Music"**, August 01, 2020 by Zijing Gao, Lichen Qiu, Peng Qi, Yan Sun:
They proposed a novel music emotion recognition model for Scratch generated music. First, we build a Scratch- generated dataset by the main melody extraction algorithm. Then, for each music, we extract their underlying features and input them to the CNN module. After that, the features learned by CNN are input to RNN to get the final classification results. In this model, the CNN module can learn the important features of music while RNN can learn the sequential features. The experimental results show that the proposed model performs better than traditional music emotion recognition models. Experiments are performed on public music datasets and homemade Scratch-generated music datasets. Compared with the use of SVM, GPC, and CNN models, CRNN has better performance on MER. Besides, the model trained using the homemade Scratch-generated music dataset has a better effect on the emotion recognition problem of Scratch-generated music.

12. **"Music Recommendation Based on Color"** October 20, 2020 by Kajornsak Kittimathaveenan, Chanathip Pongskul, Salisa Mahatanarat:
There were three stages of this study: The first stage was the preparation music library of the association between color and emotion and the association between music andemotion. Library data used for the Hue, Saturation, and Value (HSV) color model creation were: Hue to represent musical instruments, Saturation referred to tempo, and Value was key(pitch). Second stage was to create two types of GUI for color selection. The last stage was to collect data from 120 participants. The results are displayed in two formats: with numeric distributions and graphs,
- Color-to-Music Library
- GUI.

The Color-to-Music Library has low accuracy rate; which has overall accuracy only 51.11 percent.The associations of HSV model and three basic elements of music in Methods (B) should be improved. On the other hand, two types of GUI have high accuracy rates; there are 65 percent of all participants rated 4 and 5 (75 percent and above). This means that these GUIs are ready available to be used for further studies.

13. **"A system for educational and vocationalguidance: Chatbot E-Orientation"**, August 2020 by Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, Hafsa Ouchra, Oumaima Hourrane:
The DialogFlow technology, sometimes known as Speaktoit, is used to construct chatbots, Speaktoit is a Google development platform for building Natural Language Processing-based human-computer interface solutions. In our previous work we have made a classification of educational and vocational guidance questions according to Holland's theory using the neural network methods and the BERT method too. In this article, an original approach to a chatbot was introduced.

14. **"Emotion aware Smart Music Recommender System using Two Level CNN"**, 2020 by Krupa K S, Ambara G, Kartikey Rai, Sahil Choudhury:

This proposed framework is a double input CNN model that uses two approaches to detect the user's emotion through facial landmarks and semantic analysis on the interactions with a chatbot. Accuracy always increases with the dataset. Since, the dataset has greater number of images for training the happy emotion, the accuracy of the emotion turns out to be good when compared to others. There already exists widely used audio and video recommender systems like Spotify, Netflix, Gaana, YouTube etc which work based on search queries and not emotional needs of the user. Accuracy has been calculated using the Keras method. The network is trained for 60 epochs initially and the accuracy comes to be around 63% for 60 epochs. When tested for 100 epochs the accuracy came to be around 88%.

15. **"Emousic: Emotion & Activity based Music Player using Machine Learning"**, 2019 by Pranav Sarda, Sushmita Halasawade, Anuja Padmawar and Jagannath Aghav:

Here, the mood is statistically inferred from various data sources primarily: audio, image, text, and sensors. They used Random Forest Classifier for both classification and regression. Due to the large number of classes, the performance of Random forest classifier is better than Decision tree algorithms. By looking at the parameters of correctly classified instances we can check that accuracy of Random forest is more than decision tree classifier orJ48 algorithm. It consists of training dataset and test dataset where the dataset had approximately 1000 entries in training set and 100 entries in the test set. Because the experiment used a small dataset and fewer features, it might not fully represent all possible situations, which will affect the accuracy of song predictions in various scenarios.

16. **"SentiMozart: Music Generation based on Emotions"**, 2018 by Rishi Madhok, Shivali Goel and Shweta Garg:

The aim of this framework is to generate music corresponding to the emotion of the person predicted by this model. The proposed framework is divided into two models, the Image Classification Model and the Music Generation Model. The music would be generated by the latter model which is essentially a Doubly Stacked LSTM architecture. This is to be done after classification and identification of the facial expression into one of the seven major sentiment categories: Angry, Disgust, Fear, Happy, Sad, Surprise and Neutral, which would be done by using Convolutional Neural Networks (CNN). Finally, we evaluate the performance of this framework using the emotional Mean Opinion Score (MOS) which is a popular evaluation metric for audio-visual data. Even though WAV format have best quality of sound, MIDI format is one of the best options present today to generate new music as the file contains only synthesizer instructions hence the file size is hundreds of times smaller than the WAV format (Waveform Audio Format) which contains digitized sound and hence has a very large size (hundreds of megabytes).

17. **"A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks"**, October 2018 by Mohammad Nuruzzaman, Omar Khadeer Hussain:

This paper presents a survey on existing chatbots and techniques applied into it. We compared 11 most popular chatbot application systems along with functionalities and technical specifications. In the past, methods for developing chatbots have relied on hand-written rules and templates. With the rise of deep learning these models were quickly replaced by end-to-end neural networks. More specifically, Deep Neural Networks is a powerful generative based model to solve the conversational response generation problems. Based on literature review, this study made a comparison from selected papers according to method adopted. This paper also presented why current chatbot models fails to take into account when generating responses and how this affects the quality conversation. The dataset is a set of 20 QA tasks with 1000 training examples. The main aim is to fill in a gap in this research area and providing a flexible chat interface for question answering. However, the big disadvantage is that these natural responses require a great amount of learning time and data to be able to learn the vast number of possible inputs. The training will prove if the AI chatbot able to handle the more challenging issues that are normally obstacles for simpler chatbots.

18. **"Text Based Emotion Recognition: A Survey"**, 2015 by C. R. Chopade:

According to Chetan, the human-computer connection is an extremely incredible and most recent area of research in light of the fact that the human world is getting more and more digitized. This needs the computerized frameworks to emulate human behavior accurately. Chetan centered on an outline of emotion discovery from text and described the emotion detection techniques. These strategies are partitioned into the following four primary classifications: keyword-based, Lexical Affinity strategy, learning-based, and hybrid-based methodology. Restrictions of these emotion recognition strategies were presented and address the text normalization utilizing distinctive taking care of methods for both plaint text and short informing language. The application areas of textual emotion detection are:

- Sentiment Analysis
- Computer Assisted Creativity
- Text – to – Speech Generation
- Improves Human computer interaction

19. **"Approaches, Tools and Applications for Sentiment Analysis Implementation"**, Sep 2015 by Alessia D Andrea, Fernando Ferri, Patrizia Grifoni, and Tiziana Guzzo:

The machine learning approach is used for predicting the polarity of sentiments based on trained as well as test data sets. While the lexicon-based approach does not need any prior training in order to mine the data. It uses a predefined list of words, where each word is associated with a specific sentiment. The sentiment classification approaches can be classified in (i) machine learning (ii) lexicon based and (iii) hybrid approach. Using ML it has ability to adapt and create trained models for specific purposes and contexts. Lexicon-based approaches have the advantage that general knowledge sentiment lexicons have wider term coverage. The main advantages of hybrid approaches are the lexicon/learning symbiosis, the detection and measurement of sentiment at the concept level and the lesser sensitivity to changes in topic domain,

V. REQUIREMENT SPECIFICATION

Hardware Requirements

Hardware requirements refer to the minimum and recommended specifications of computer components necessary to run a particular software or application efficiently. These specifications typically include details such as the required processor speed, amount of RAM, available storage space, and any specific graphics capabilities needed. Following are the different hardware requirements:

- Processor: Intel i5
- RAM: 8 GB
- Hard disk space: 1TB

Software Requirements

Software requirements refer to the specifications and functionalities that a software application must meet to satisfy the needs of users and stakeholders. These requirements outline the features, performance expectations, compatibility criteria, and other criteria that guide the design, development, and testing of the software. They serve as a foundation for communication between developers and stakeholders, ensuring that the final product meets the intended objectives and functions effectively. Following are the different software requirements:

- Programming Languages: Python
- MLFrameworks: TensorFlow, PyTorch
- NLPLibraries: NLTK (Natural Language Toolkit), spaCy
- Computer Vision Libraries: OpenCVAudio
- Framework: Django
- DBMS: MySQL
- IDE: PyCharm, Jupyter Notebooks

I. PROPOSED WORK

"EmoTune: Emotion and Gender-Aware Music Generation Chatbot" is an innovative AI-driven project that redefines the user's musical journey. This chatbot seamlessly processes real-time text, images, audio, and video inputs, responding dynamically to user emotions and gender preferences. With features like classic song generation for aged users and multilingual lyrics capabilities, EmoTune celebrates cultural diversity and fosters cross-generational music appreciation. The project's commitment to ethical AI, user-friendly interfaces, and compatibility with popular development environments ensures a responsible and accessible musical companion. EmoTune envisions a future scope that includes advanced emotion recognition models, integration with smart devices, and immersive controls through voice and gesture commands, promising an unparalleled and personalized musical experience for users worldwide.

Here we propose Emotune, a cutting-edge mobile application designed to enhance the music discovery experience by tailoring song recommendations based on the user's emotional state. This innovative app utilizes facial analysis through the mobile camera to discern the user's emotional condition, subsequently offering song suggestions aligned with the detected mood. The emotional spectrum is intelligently categorized into four primary types: neutral, happy, sad, and angry. In today's world, when we talk to computers or use interactive systems, it's hard for them to understand how we feel. It's important to have a system that can quickly recognize our emotions in real-time, especially for strong emotions like sadness, anger, or depression. People sometimes struggle to share their personal feelings, making it tough to seek help for mental health. That's why many hope that machines can not only understand but also respond appropriately to our emotions.

Natural Language Processing (NLP) emerges as a pivotal technology in achieving this level of understanding. Within the realm of NLP, Emotion Detection and Recognition from Text is a rapidly evolving field intricately linked to sentiment analysis. Detecting emotions from textual data presents a complex challenge due to the contextual nuances inherent in language. For example, the phrase "Shut up!" conveys anger without explicitly using the term. In this context, our application adopts advanced Deep Learning techniques within the Chatbot module. It employs a Long Short-Term Memory (LSTM) recurrent neural network to facilitate effective message categorization and response generation. EmoTune's approach is multifaceted, aiming to navigate the intricate interplay between emotions, music, and technology. The goal is to provide users with a personalized and emotionally resonant music experience. By combining cutting-edge facial analysis, NLP, and deep learning, Emotune seeks to redefine how users interact with and experience music, creating a seamless bridge between emotional states and musical preferences.

EmoTune, the Emotion and Gender-Aware Chatbot, is like a carefully designed plan showing how different parts of the system work together. Imagine it as a map that explains how this smart chatbot understands your feelings and figures out if you're a guy or a girl based on what you say or type. it listens to what you say or type and uses special tricks to understand how you're feeling and to guess your gender. Then, it uses these clues to decide how it should talk back to you. All these actions happen in a step-by-step way. This plan not only shows how these parts connect but also how they can handle more people talking to the chatbot at once, keep things safe and private, and even learn to get better over time by paying attention to how people react to it. Essentially, it's a detailed plan that makes sure the chatbot can understand and respond to you in a smart and personalized way.

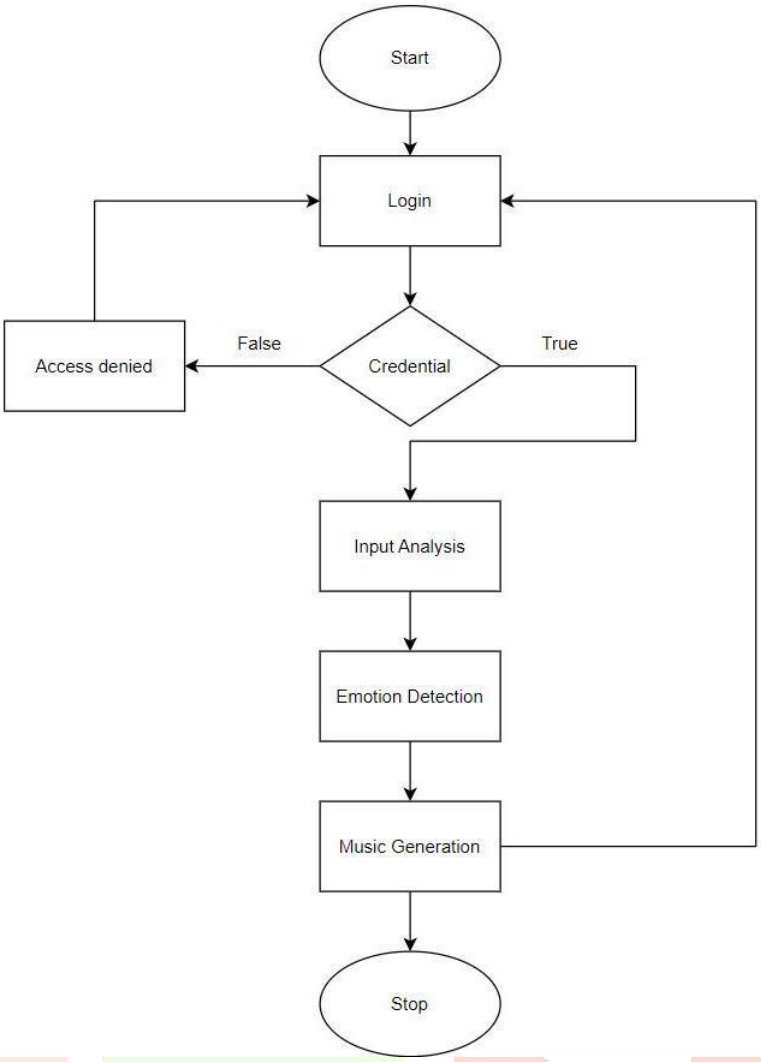

Fig 1. Flowchart of the project

Above fig 1 shows the flow chart of EmoTune: Emotion and Gender Aware Music Generation Chatbot. First, the user has to login. If the entered credentials are proper, chatbot interface will be displayed. Else, access will be denied. Once the chatbot interface is opened, the user can interact with the chatbot, text is analyzed and then the emotion is detected. Later, according to the emotions detected, it will generate the songs. These generated songs are played to the user.

II. SYSTEM ARCHITETURE

The software needs the architectural design to represents the design of software. IEEE defines architectural design as "the process of defining a collection of hardware and software components and their interfaces to establish the framework for the development of a computer system." The software that is built for computer-based systems can exhibit one of these many architectural styles. Architectural design is a creative process where you try to establish a system organization that will satisfy the functional and non-functional system requirements. Because it is a creative process the activities within the process differ radically depending on the type of system being developed, the background and experience of the system architect, and the system requirements for the system. An architecture diagram is a network map used to describe the general structure of a software program as well as the interactions, restrictions and limits between elements. It involves creating a plan that outlines the various components, modules, and subsystems that make up the system, as well as how they interact with one another. Fig 2 shows the System Architectural of EmoTune: Emotion and Gender-Aware Chatbot. EmoTune is an AI-powered Music Generation Chatbot that's changing how we experience personalized music. This smart system can understand and respond to what you say, show, or play in real-time. It tailors its responses based on your emotions and the type of music you like. EmoTune cares about your privacy and is easy to use, fitting well with different devices. It aims to be a responsible and inclusive music companion for everyone. It's not just about changing how we listen to music now; it's also thinking about the future of music, where it's all about

your unique musical journey. Description of the architecture diagram of this system including the steps from user profiling to final output:

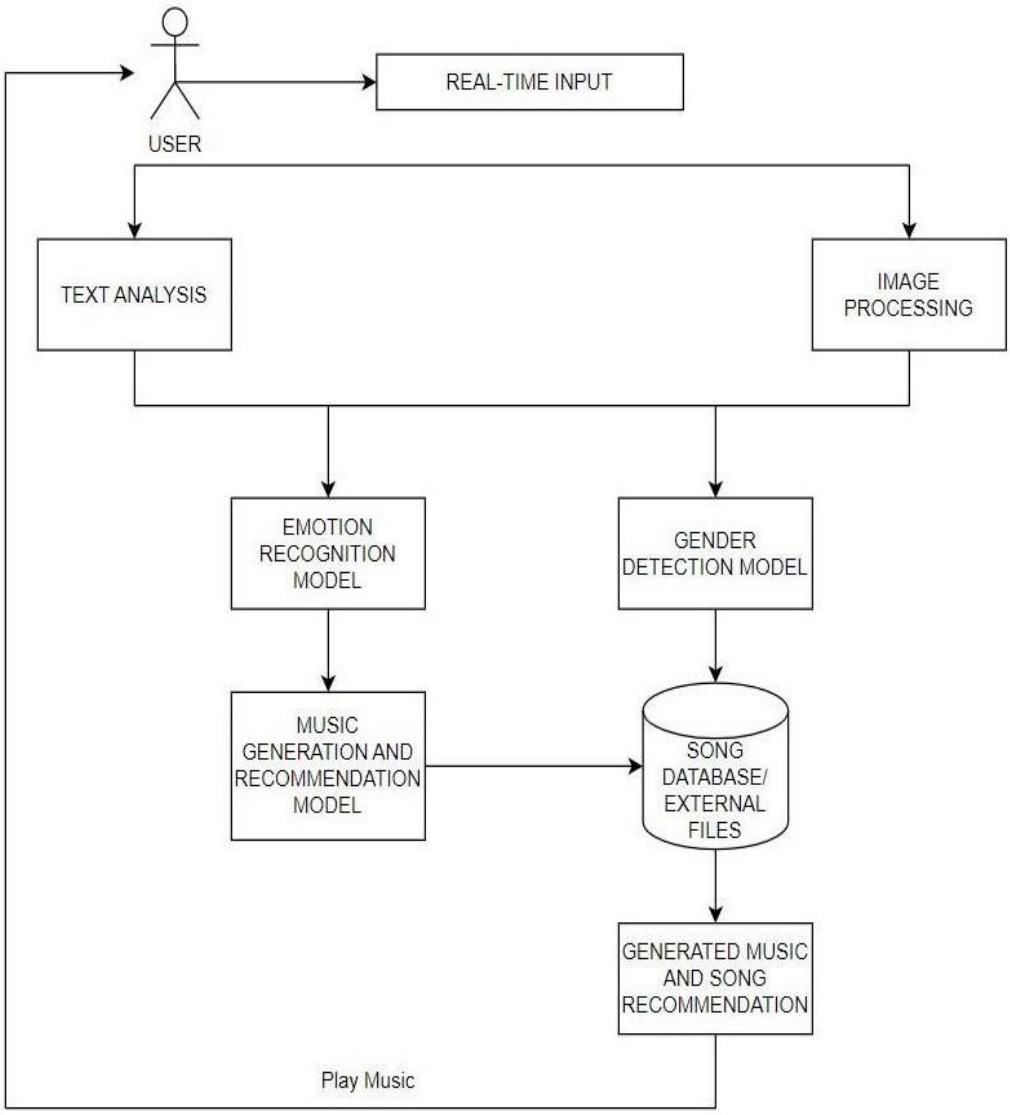

Fig 2 System Architecture of Emotune

Above Fig 2 shows the proposed system of EmoTune: Emotion and Gender Aware Chatbot. First, User will have to give the Input. Input can be in the form of text, audio or images. If the text is given as input, i.e. it will have to interact with the chatbot and then text analysis and NLP models are performed based on the chatbot interaction taken place. If user has given input in the form of image, then image processing takes place and if the user has given input in the form of audio, then audio analysis and speech emotion models are performed. After the emotion is recognized, it will take the information about the music based on the emotion from the database. In the same way, Gender detection takes place and it will take information from the database about the music based on gender. Then, it will generate music to the user. Finally, it plays the music to the user.

I. CONCLUSION

In conclusion, "EmoTune: Emotion and Gender-Aware Music Generation Chatbot" stands as a transformative embodiment of the evolving synergy between artificial intelligence and the realm of musical experiences. This innovative project has successfully woven together real-time multimodal input processing, advanced emotion recognition, and gender sensitive music generation to craft a deeply personalized and inclusive musical companion. EmoTune not only responds dynamically to user emotions and preferences but also pays homage to classic songs, recognizing aged users and generating nostalgic compositions. Its multilingual lyrics generation fosters cross-cultural storytelling, while the commitment to

ethical considerations, user-friendly interfaces, and compatibility with popular development environments underscores its dedication to responsible and accessible AI. Looking ahead, the future scope envisions enhancements such as advanced emotion recognition models, expanded language support, and integration with smart devices and virtual reality environments. With potential features like voice and gesture commands for playback control, the project aims to create an immersive and intuitive interaction experience. EmoTune, therefore, transcends the boundaries of traditional music creation, offering a unique fusion of technology and emotion. As a versatile and innovative musical companion, it not only connects users with music on a profound level but also paves the way for continued exploration and creativity in the dynamic intersection of artificial intelligence and musical expression. EmoTune, with its groundbreaking features and future possibilities, signifies a paradigm shift in how we engage with and derive emotional resonance from the vast world of music.

II. ACKNOWLEDGEMENT

We would like to express our gratitude and thanks to Dr. Shantharama Rai C, Principal, at A J Institute of Engineering and Technology for providing us a good project guide during a lot of consultation. We also take this opportunity to convey a deep feeling of gratitude to Dr. Suresha D, HOD of Information Science and Engineering, with his warm support, valuable information and guidance. Thank you for your co-operation sir. We take this opportunity to extend our deepest gratitude and deepest greetings to our Guide Prof. Archana Priyadarshini Rao for presenting her work ethic and her exemplary guidance, vigilance, and ongoing encouragement throughout the course We would like to thank all the people for their direct and indirect assistance in completing our work.

REFERENCES

[1] “Chatbot Song Recommendation System”, International Journal of Engineering Research & Technology (IJERT) - 2023 by Anusha, Dr. Srinivasan V.
[2] “Chatbot with Song Recommendation based on Emotion”, 2022 by Simran Chaudhari, Hrucha Malusare, Aadil Adheesh, Mokshada Bhadavalkar, Asma Shaikh.
[3] "Mood Based Music Playlist Generator Using Convolutional Neural NetworkProf", International Journal of Engineering Research & Technology (IJERT) - 2022 by Jaychand Upadhyay, Sharan Shetty, Vaibhav Murari, Jarvis Trinidade.
[4] "Alexis: A Voicebased Chatbot using Natural Language Processing", International Journal of Engineering Research & Technology (IJERT) - 2022 by Meet Popat, Aayush Doshi, Yashraj Rai.
[5] "Emotion Recognition Using Chatbot System", Research Gate - 2022 by Shraddha Pophale, Hetal Gandhi and Anil Kumar Gupta.
[6] "A Machine Learning Based Chatbot Song Recommender System", 2021 by Amaan Shaikh, Bhushan Patil, Tejas Sonawane.
[7] “Music Recommender System Using ChatBot”, 2021 by Shivam Sakore, Pratik Jagdale, Mansi Borawake, Ankita Khandalkar.
[8] "A Novel Music Emotion Recognition Model Using Neural Network Technology", Frontiers in Psychology -28 Sep 2021 by Jing Yang.
[9] "Music recommendation system based on facial emotion recognition", March 2020 by Deny John Samuel, B. Perumal, Muthukumaran Elangovan.
[10] “Conversational and Image Recognition Chatbot”, Stanford CS224N - 2020 by Shengyang Su.
[11] “A Novel Music Emotion Recognition Model for Scratch generated Music” IEEE Xplore August 01, 2020 by Zijing Gao,
Lichen Qiu, Peng Qi, Yan Sun
[12] “Music Recommendation Based on Color” IEEE Explore - October 20, 2020 by Kajornsak Kittimathaveenan, Chanathip Pongskul, Salisa Mahatanarat.
[13] “A system for educational and vocationalguidance: Chatbot E-Orientation”, August 2020 by Omar Zahour, El Habib Benlahmar, Ahmed Eddaoui, Hafsa Ouchra, Oumaima Hourrane.
[14] “ Emotion aware Smart Music Recommender System using Two Level CNN” IEEE Xplore Part Number: CFP20P17 - ART; ISBN: 978-1-7281-5821-1 by Krupa K S, Ambara G, Kartikey Rai, Sahil Choudhury.
[15] "Emousic: Emotion & Activity based Music Player using Machine Learning", Research Gate - 2019 by Pranav Sarda, Sushmita Halasawade, Anuja Padmawar and Jagannath Aghav.
[16] "SentiMozart: Music Generation based on Emotions", 10th International Conference on Agents and Artificial Intelligence (ICAART) - 2018 by Rishi Madhok, Shivali Goel and Shweta Garg.

[17] "A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks" IEEE Xplore October 2018 by Mohammad Nuruzzaman, Omar Khadeer Hussain.
[18] "Text Based Emotion Recognition: A Survey", 2015 by C. R. Chopade.
[19] "Approaches, Tools and Applications for Sentiment Analysis Implementation", Sep 2015 by Alessia D Andrea, Fernando Ferri, Patrizia Grifoni, and Tiziana Guzzo.
[20] "Geometric Approach for Human Emotion Recognition using Facial Expression", International Journal of Computer Applications, 2015by Bavkar, Sandeep, Rangole, Jyoti, Deshmukh.
[21] "Emotion Based Music Player", IEEE 2019 conference by Krittrin Chankuptarat, etal.
[22] "A Comparative Study of Music Recommendation Systems" 2018 IEEE by Ashish Patel , Dr. Rajesh Wadhvani.
[23] "A Survey on Chatbot Implementation in Customer Service Industry through Deep Neural Networks" IEEE Xplore October 2018 by Mohammad Nuruzzaman, Omar Khadeer Hussain.
[24] "MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction" CIKM '19, November 3– 7, 2019 by Yucheng Jin, Wanling Cai, Li Chen, Nyi Nyi Htun, Katrien Verbert.
[25] Competence-Based Song Recommendation: Matching Songs to One's SingingSkill" IEEE Transactions on Multimedia,VOL. 17, NO. 3, March 2015 by Kuang Mao, Lidan Shou,Ju Fan, Gang Chen and Mohan S. Kankanhalli.