

SONG RECOMMENDATION BASED ON VOICE TONE ANALYSIS

¹ A.Rajesh, Assistant Professor, Department of CSE, Chalapathi Institute of Technology, Guntur.

² Thatiparthi Pavani, B.Tech, Department of CSE, Chalapathi Institute of Technology, Guntur.

³ Ravuri Venkata Sai Naveen, B.Tech, Department of CSE, Chalapathi Institute of Technology, Guntur.

⁴ Puchakayala Manikanta, B.Tech, Department of CSE, Chalapathi Institute of Technology, Guntur

⁵ Pathuri Sandeep, B.Tech, Department of CSE, Chalapathi Institute of Technology, Guntur.

ABSTRACT

Music suggestions based on elements including user listening history, music genre, etc. have become more popular in recent years. The user's emotional state is not considered by conventional music recommendation algorithms; therefore, they might not offer suggestions that are appropriate for their present mood. This study introduces a song recommendation system that uses artificial intelligence and machine learning to make individualized music suggestions based on the emotional state of the user. The system extracts feature with the help of MFCC to analyze the tone of the user. The suggested method incorporates deep learning models such as Artificial Neural Networks that provide better accuracy to train the model. The major challenge in creating such a system is to successfully determine the data for the recommendation process by accurately and consistently detecting the user's emotional state from speech. The proposed system offers new pathways for research in the fields of artificial intelligence and music recommendation and has the potential to alter the way people listen to music.

I.INTRODUCTION

1.1 Introduction

Singing skills of each singer helps in choosing accurate songs for them. These skills are determined by using vocal competence. Singers with different singing skills give better performance only if song is suitable to their skills. Many times, singers are upset due to their performance. The reason behind it is sometimes because of wrong preferred song by singer than the singing ability of singer. E.g. it is difficult task for a girl having soft voice to sing a song which requires strong voice. To achieve a good performance, songs are chosen strongly with singer's vocal competence.

Matching a song to one's vocal ability is the major task in song recommendation. Experienced singing

teachers listens the voice of singer to find point in a voice. This point then used to choose suitable song matching with singer's voice. Typically, they choose tough songs for singers which differentiate singers from each other. They recommend songs which gives assurance of good singing performance.

Radio is one of the mass media closely related to society's needs, which can provide various kinds of information, entertainment, and education. Radio as an effective mass media in disseminating audio-based information, various kinds of information can be conveyed in clear audio and in an easily understood language by the public [1]. During a radio broadcast, there is an interaction between broadcasters and radio listeners. Radio listeners can request song and give greetings to the broadcaster. Then, broadcasters will later play the requested songs. When this research was conducted, there was a problem in the song request process. The first is that there is no place to request a song. The second problem is from the listeners' side that the listeners do not know how to request the appropriate song.

The absence of song list makes listeners not know what songs are suitable for them. The problem is the difficulty of broadcasters in monitoring song requests from listeners in order to make a system can accommodate all song requests by listeners. The way that can be done to request songs is by online telephone [2]. Song request using the telephone is done by utilizing the Voice over Internet Protocol (VoIP) network technology [3]. Technically, VoIP technology converts analog voice into digital data packets, which are later transferred over the Internet Protocol (IP). In this study, the VoIP client used is Sinch [4]. Many greetings from listeners that come through several social media accounts make broadcasters have to work twice. Thus, a song request system is designed by saving messages to the database, which later will be retrieved and displayed to the broadcaster. The concept of song request will

be the same as the queuing process to be more in order [5].

1.2 Problem statement

Music has become an integral part of our day-to-day life. With the advancement in technology, lots of businesses are nowadays using recommendation systems for their benefit like in e-commerce, selling books, movies, and video recommendations. This helps businesses to get monetary benefits and the users to get better services. One of the best ways to suggest songs would be according to the people's mood. In this project, we aim to interpret the user's mood at a particular time using what they speak and then suggest songs based on the analysis done by the system. The best thing about music is that nothing can be more relaxing than a pleasant melody. Due to all these things, we choose to do this Song Recommendation system according to the mood of the user.

1.3 Objectives

To find favorite songs among massive songs has become a difficult problem. The song recommendation algorithm makes personalized recommendations by analyzing user's historical behavior, which can reduce user's information fatigue and improve the user experience. This paper studies a personalized song recommendation algorithm based on vocal features. The specific work includes three parts. Firstly, the spectrum feature extraction and observe feature extraction of songs. The spectrum includes three types of features: time domain, frequency domain, and amplitude, which implicitly describe the rhythm, notes, and high-pitched or soothing properties of songs. Furthermore, automatic note recognition methods are explored as explicit classification features.

II.LITERATURE SURVEY

Song Recommendation Based on Vocal Competence Ms. Samruddhi S. Doshi¹, Mr. Dinesh B. Hanchate² Student, Computer Engineering, VPKBIET, Baramati, India 1 Computer Engineering, VPKBIET, Baramati, India 2

Singing skills of each singer helps in choosing accurate songs for them. These skills are determined by using vocal competence. Matching a song to singer's vocal competence is the major and difficult task in song recommendation. There are other song recommendation systems, which works on recommending songs of listener's interest. These traditional approaches did not consider the singer's ability of singing which leads to poor singing performance by singer. This paper gives the solution using song recommendation framework. This framework uses the term singer profile, which constitutes singer's vocal competence in terms of features of voice like pitch, intensity and also the voice quality of the singer. The paper then gives technique to acquire singer profile and song profiles are stored in database. Then learning to- rank scheme is used to rank the recommended song profile for each singer profile.

Personalized song recommendation method based on voice timbre: The invention provides a personalized song recommendation method based on voice timbre, relates to a network platform singing song recommendation method, and solves the problem that a collaborative filtering method is not suitable in the case and a speaker recognition correlation method cannot be used directly, and thus large influence is caused on the recommendation result. The method comprises the following steps: adopting accompaniment and voice separation; 2) extracting acoustic features MFCC and LPCC capable of representing user timbre features from the voice data; 3) recommending personalized required songs according to the acoustic features extracted in the step 2); and 4) carrying out expandability analysis. The method improves user experience in a network singing system and has a very good application prospect.

Song recommendation using speech emotion recognition : Music has become an integral part of our day-to-day life. With the advancement in technology, lots of businesses are nowadays using recommendation systems for their benefit like in e-commerce, selling books, movies, and video recommendations. This helps businesses to get monetary benefits and the users to get better services. One of the best ways to suggest songs would be

according to the people's mood. In this project, we aim to interpret the user's mood at a particular time using what they speak and then suggest songs based on the analysis done by the system. The best thing about music is that nothing can be more relaxing than a pleasant melody. Due to all these things, we choose to do this Song Recommendation system according to the mood of the user.

Face Detection and Facial Expression Recognition System

AUTHOR: Anagha S.Dhavalikar, and Dr. R. K. Kulkarni,

A human-computer interaction system for an automatic face recognition or facial expression recognition has attracted increasing attention from researchers in psychology, computer science, linguistics, neuroscience, and related disciplines. In this paper, an Automatic Facial Expression Recognition System (AFERS) has been proposed. The proposed method has three stages: (a) face detection, (b) feature extraction and (c) facial expression recognition. The first phase of face detection involves skin color detection using YCbCr color model, lighting compensation for getting uniformity on face and morphological operations for retaining the required face portion. The output of the first phase is used for extracting facial features like eyes, nose, and mouth using AAM (Active Appearance Model) method. The third stage, automatic facial expression recognition, involves simple Euclidean Distance method. In this method, the Euclidean distance between the feature points of the training images and that of the query image is compared. Based on minimum Euclidean distance, output image expression is decided. True recognition rate for this method is around 90% - 95%. Further modification of this method is done using Artificial Neuro-Fuzzy Inference System (ANFIS). This non-linear recognition system gives recognition rate of around 100% which is acceptable compared to other methods.

III.SYSTEM ANALYSIS

EXISTING SYSTEM :

Chat-bots can only handle basic questions: Chat-bots are still a basic Artificial Intelligence technology and so they can only answer the basic questions of customers and provide general information that is already available to them. They cannot solve complicated queries or answer out of script questions and companies need to have human customer service employees that can manage these for them. However, this is changing with time and currently, more and more advanced chat-bots are entering the market. Chat-bots are difficult to create: Chat-bots are created using Natural Language Processing which is extremely popular for customer support applications. Natural Language Processing is a part of Machine Learning which can be used to interact with the users in textual form and solve their queries. However, 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.

DISADVANTAGES OF EXISTING SYSTEM :

- 1) Less accuracy
- 2)low Efficiency

PROPOSED SYSTEM :

In this project we have designed deep learning based emotion detection model from voice pitch and based on user pitch emotion Chatbot will suggest songs. Chatbot ask user to record his voice and then analyse user pitch to detect emotion and based on emotion songs will be recommended.

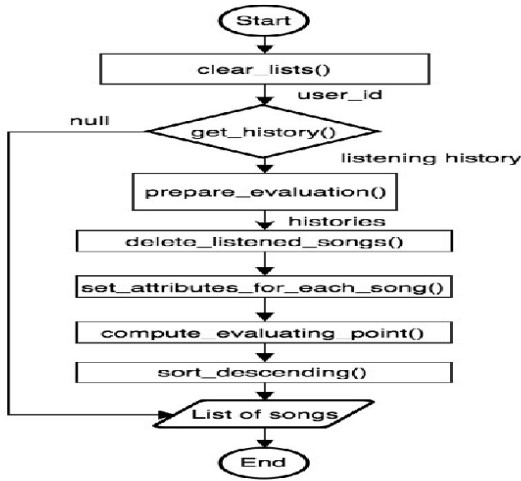
Propose model can identify 8 different emotions such as Happy, Calm, neutral which we are considering as Happy and suggesting JOSH songs. This model will predict other emotions like Sad, Angry, feared, disgust and surprise which are considering user as in SAD mood and suggesting HAPPY songs.

On demand also songs will be suggested where user has to type his feeling as Happy or Josh and based on that songs will be recommended.

ADVANTAGES OF PROPOSED SYSTEM :

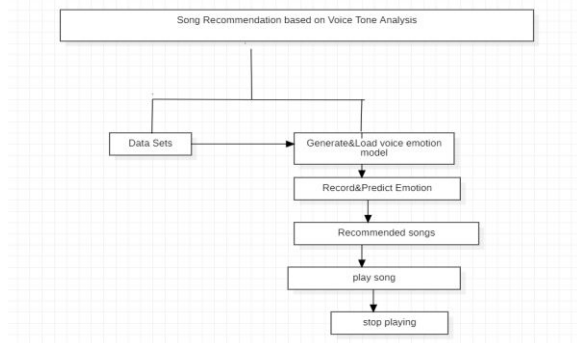
- 1) High accuracy
- 2)High efficiency

IV.SYSTEM ARCHITECTURE



Voice Tone is the frontend of our project where our users interact, We will be using the API functionality in this module, Voice Tone we collect the user input using a get parameter and this input is sent to our second module Emotional Tone Analyzer. We will be using the POST request to get the chat response we will be providing the post request with two parameters context and emotion. The context is the user input, and the emotion is analyzed from the emotion tone analyzer we must also create a mapping for the emotion received from the emotional tone analyser since CakeChat can recognize only one of the following emotions {'neutral', 'anger', 'joy', 'fear', 'sadness'}, emotion to condition the response on, optional param if not specified, 'neutral' is used. Once the API request is sent, we get a JSON data with the response this JSON data is extracted and is used to display the Voice Tone response in the Tkinter.

**DETAILED DESCRIPTION
BLOCK DIGRAM**



This is the detailed representation of our project

Modules Description

In this paper author describing concept to automate government services with Artificial Intelligence technology such as Deep Learning algorithm called Convolution Neural Networks (CNN). Government can introduce new schemes on internet and peoples can read news and notifications of such schemes and then peoples can write opinion about such schemes and this opinions can help government in taking better decisions. To detect public opinions about schemes automatically we need to have software like human brains which can easily understand the opinion which peoples are writing is in favour of positive or negative.

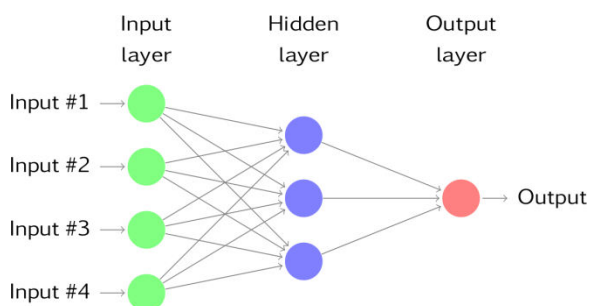
To build such automated opinion detection author is suggesting to build CNN model which can work like human brains. This CNN model can be generated for any services and we can make it to work like automated decision making without any human interactions. To suggest this technique author already describing concept to implement multiple models in which one model can detect or recognize human hand written digits and second model can detect sentiment from text sentences which can be given by human about government schemes. In our extension model we added another model which can detect sentiment from person face image. Person face expressions can describe sentiments better than words or sentences. So our extension work can predict sentiments from person face images.

To demonstrate how to build a convolutional neural network based image classifier, we shall build a 6 layer neural network that will identify and separate one image from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing image classification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world convolutional neural network using TENSORFLOW.

Neural Networks are essentially mathematical models to solve an optimization problem. They are made of neurons, the basic computation unit of neural networks. A neuron takes an input (say x), do some

computation on it (say: multiply it with a variable w and adds another variable b) to produce a value (say; $z = wx + b$). This value is passed to a non-linear function called activation function (f) to produce the final output (activation) of a neuron. There are many kinds of activation functions. One of the popular activation function is Sigmoid. The neuron which uses sigmoid function as an activation function will be called sigmoid neuron. Depending on the activation functions, neurons are named and there are many kinds of them like RELU, TanH.

If you stack neurons in a single line, it's called a layer; which is the next building block of neural networks. See below image with layers



To predict image class multiple layers operate on each other to get best match layer and this process continues till no more improvement left.

V.SCREEN SHOTS

Song Recommendation based on Voice Tone Analysis

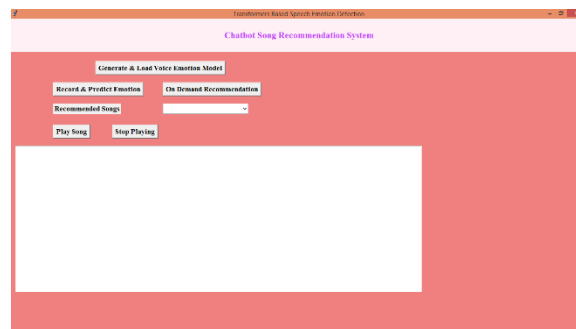
In this project we have designed deep learning based emotion detection model from voice pitch and based on user pitch emotion Voice Tone will suggest songs. Voice Tone ask user to record his voice and then analyse user pitch to detect emotion and based on emotion songs will be recommended.

Propose model can identify 8 different emotions such as Happy, Calm, neutral which we are considering as Happy and suggesting JOSH songs. This model will predict other emotions like Sad, Angry, feared, disgust and surprise which are considering user as in SAD mood and suggesting HAPPY songs.

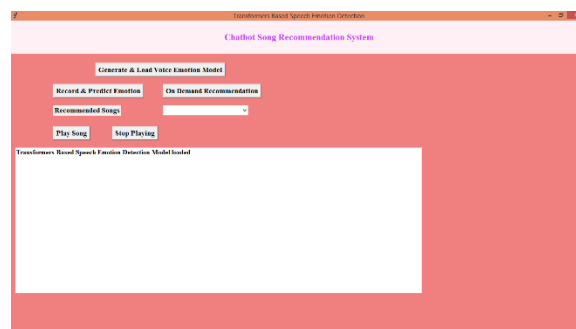
On demand also songs will be suggested where user has to type his feeling as Happy or Josh and based on that songs will be recommended.

SCREEN SHOTS

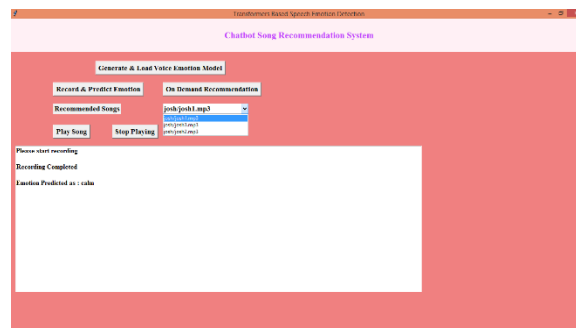
To run project double click on 'run.bat' file to get below screen



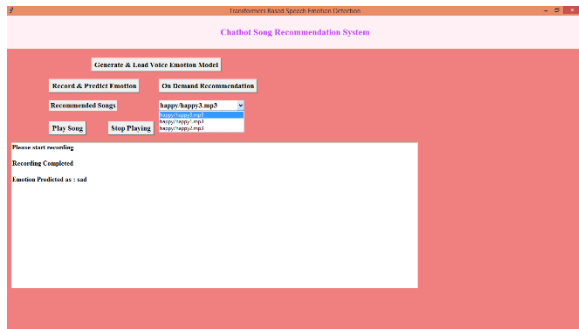
In above screen click on 'Generate & Load Voice Emotion' model button to load model and get below output



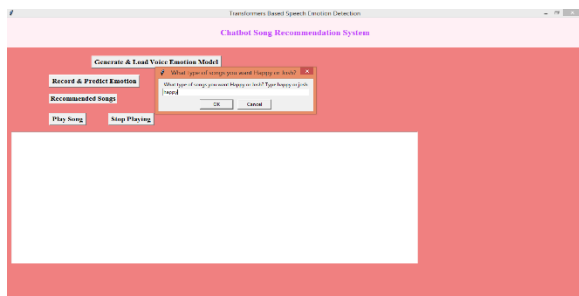
In above screen model is loaded and now click on 'Record & Predict Emotion' button to start recorder and then you need to speak something and once model get sufficient voice data then it will predict emotion based on pitch and give below output



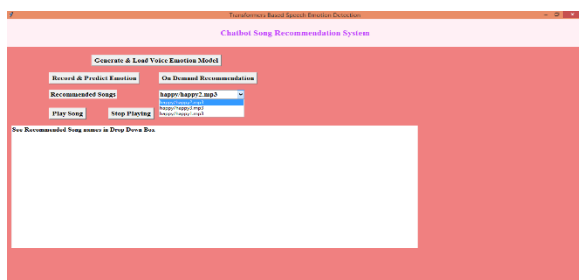
In above screen in text area we can see predicted emotion from pitch is calm and based on predicted emotion we can see list of recommended JOSH songs in drop down list and user can select desired song and click on ‘Play Song’ button to play the song or “Stop Playing” button to stop player. Now in below screen will try another voice from different pitch



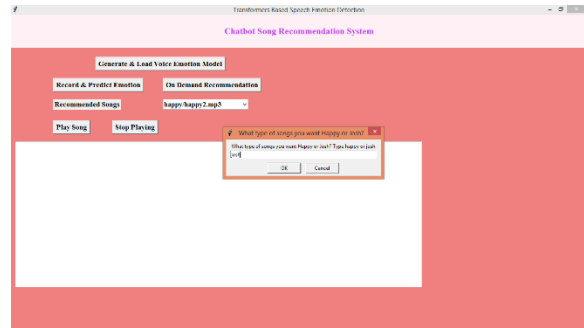
In above screen from voice pitch emotion detected as ‘sad’ and then displaying ‘HAPPY’ songs in dropdown list. Similarly record your voice and get recommended songs from Voice Tone . Now click on ‘On Demand Recommendation’ button and user can type words based on his condition and get recommendation



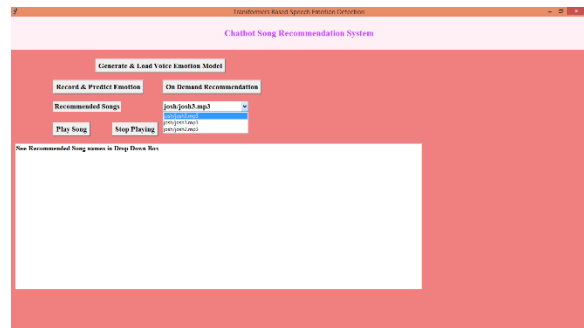
In above screen user is asking for happy songs and below is the recommendation



In above screen in drop down box user is got list of happy songs and below is another example



In above screen user is asking for JOSH songs and below are the recommendation



In above screen we got all JOSH songs as recommendation.

Note: if you want you can put new songs inside ‘josh or happy’ folders to get more recommendation. In same application Voice Tone allow you to play and stop song playing.

VI. CONCLUSION

In this paper, SentiSpotMusic: a music recommendation framework, in regard to the sentiments has been proposed using Tableau Dashboard and Spotify dataset. The exploratory data & sentiment analysis revealed a number of groundbreaking revelations. When an artist is already popular, the new songs by them have a much higher chance of gaining popularity. However, in general, to increase the odds, the song has to have high energy and danceability. Non acoustic Songs that are loud tend to garner higher popularity than others. The songs that were released after 2006 are quite popular till date. The PCC results weren’t significant. Hence, there is no need to complicate things. Keeping the track length between 3-5 minutes is observed to be a commonality among the songs. Those songs that are

recorded in a studio (not live) with lesser speechiness have a fair chance of getting popular. Lastly, high valence, i.e. cheerful ones, tend to gain more popularity

FEATURE ENHANCEMENT

FUTURE WORK The future work for the project includes building the recommendation system as per the proposed methodology on the basis of key factors and results derived from the sentiment analysis of personalized Spotify playlist. The key factors would involve liveliness, danceability, loudness, valence and such. The results of sentiment analysis would predict the amount of positive, neutral and negative sentiment in the songs involved.

References

- [1] K. K. Kumar, S. G. B. Kumar, S. G. R. Rao and S. S. J. Sydulu, "Safe and high secured ranked keyword searchover an outsourced cloud data," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 20-25, doi: 10.1109/ICICI.2017.8365348.
- [2] A. S. B. Nugroho, I. Wardiah, and M. K. Ilmi, "Aplikasi Radio Daring Pemerintah Kabupaten Hulu Sungai Tengah Berbasis Web," *Poros Tek.*, vol. 10, no. 1, 2018.
- [3] M. Exsan and U. Fadlilah, "Pembangunan Infrastruktur Voice Over Internet Protocol di Organisasi Perangkat Daerah Boyolali menggunakan Server Elastix," *Emit. J. Tek. Elektro*, vol. 17, no. 2, 2017.
- [4] K. K. . Kommineni and A. . Prasad, "A Review on Privacy and Security Improvement Mechanisms in MANETs", *Int J Intell Syst Appl Eng*, vol. 12, no. 2, pp. 90–99, Dec. 2023..
- [5] S. B. Aziz, T. A. Riza, and R. Tulloh, "PERANCANGAN DAN IMPLEMENTASI APLIKASI SISTEM ANTRIAN UNTUK PASIEN PADA DOKTER UMUM BERBASIS ANDROID DAN SMS GATEWAY," *J. Elektro dan Telekomun. Terap.*, vol. 2, no. 1, 2016.
- [6] J. M. Garcia-Garcia, V. M. R. Penichet, and M. D. Lozano, "Emotion detection: A technology review," in *ACM International Conference Proceeding Series*, 2017, vol. Part F131194.
- [7] M. Magdin, T. Sulka, J. Tomanová, and M. Vozár, "Voice Analysis Using PRAAT Software and Classification of User Emotional State," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 5, no. 6, 2019.
- [8] G. Li and J. Zhang, "Music personalized recommendation system based on improved KNN algorithm," in *Proceedings of 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference, IAEAC 2018*, 2018.
- [9] J. Sarah, "Pemanfaatan Api Spotify Dan Twitter Dalam Merekomendasikan Lagu Berdasarkan Mood Pada Tweet Dan Aktivitas Pengguna Melalui Sensor Accelerometer Berbasis Android," *Universitas Komputer Indonesia*, 2018.
- [10] Kiran Kumar Kommineni, Ratna Babu Pilli, K. Tejaswi, P. Venkata Siva, Attention-based Bayesian inferential imagery captioning maker, *Materials Today: Proceedings*, 2023, , ISSN 2214-7853, <https://doi.org/10.1016/j.matpr.2023.05.231>.