

EMOTION DETECTION USING VOICE AND SONG SUGGESTION

¹Priyanka Kale, ²Sonal Gogavle, ³Shital Kalantre, ⁴Prof. J.A Gaikwad,

^{1, 2, 3} B.E Computer Engineering, ⁴ Assistant Professor

^{1, 2, 3, 4} Department of Computer Engineering / Datta Meghe College of Engineering, Airoli, Navi Mumbai

^{1, 2, 3, 4} Mumbai University, Mumbai, India

Email - ¹priyankakale708@gmail.com, ²sonalgogavle@gmail.com, ³shitalkalantre11as@gmail.com

⁴jag.cm.dmce@gmail.com

Abstract: Generally speaking, voice and expression are natural and clear ways for human beings to interact with the world. Voice is a main aspect of verbal communication. Emotion detection using an automated voice detector would automatically convert the voice to text to know what the user's mood is, what the user's requirements are. On the basis of the language, the program tests the similarity of language terms with the terms in the code. Check the algorithm for the match. If the match algorithm finds matches, the program would recommend and display the songs stored in the device database to the user's requirements. The key idea of this project is to recommend a song that suits the mood of the consumer.

Key Words: Voice, language, Similarity, Mood, Emotion detection.

1. INTRODUCTION:

1.1. Emotion Detection:

Human Voice Recognition is one of the most efficient and strong functions in social communication. Voice is an essential process of communication that allows the ability to communicate with other people. Voice and speech are usually common and clear ways for human beings to interact with the world. Voice is the main aspect of verbal communication. Human voice pitch often varies from emotion to emotion. This program explains how the human voice can be used and interpreted to recommend songs based on user speech. The machine considers the user's mood based on his / her speech as an input to the device. The machine processes the data and translates it to text. The Cloud Speech API allows the conversion of voice to text. The Answer can take some time to get back.

1.2. Song Suggestion:

When the text processing is completed, the program compares the language being interpreted with the words that have been stored in the program database. If a match is identified with the interpreted text of the user, the program will recommend a matched song and show a list of songs. As if the interpreted text contains terms linked to emotions that are sad, joyful, anxious, nervous, the algorithm can look for a word or phrase in a scheme that suits all of them in particular. Before that, it fetches the songs that are in the matched group and advises the user by showing the list of songs.

1.3. Traditional Approach of Voice and Song:

Now we can see that there are a range of Android and Web applications that offer music services. Such applications are based on making a playlist of similar songs and often searching for that specific song by the name of the album, the name of the artists, etc. Such a conventional method demands that things be performed manually. That is the biggest downside of current programs. Where this method overcomes the limitations of the conventional approach by using a voice instead of manually creating a playlist.

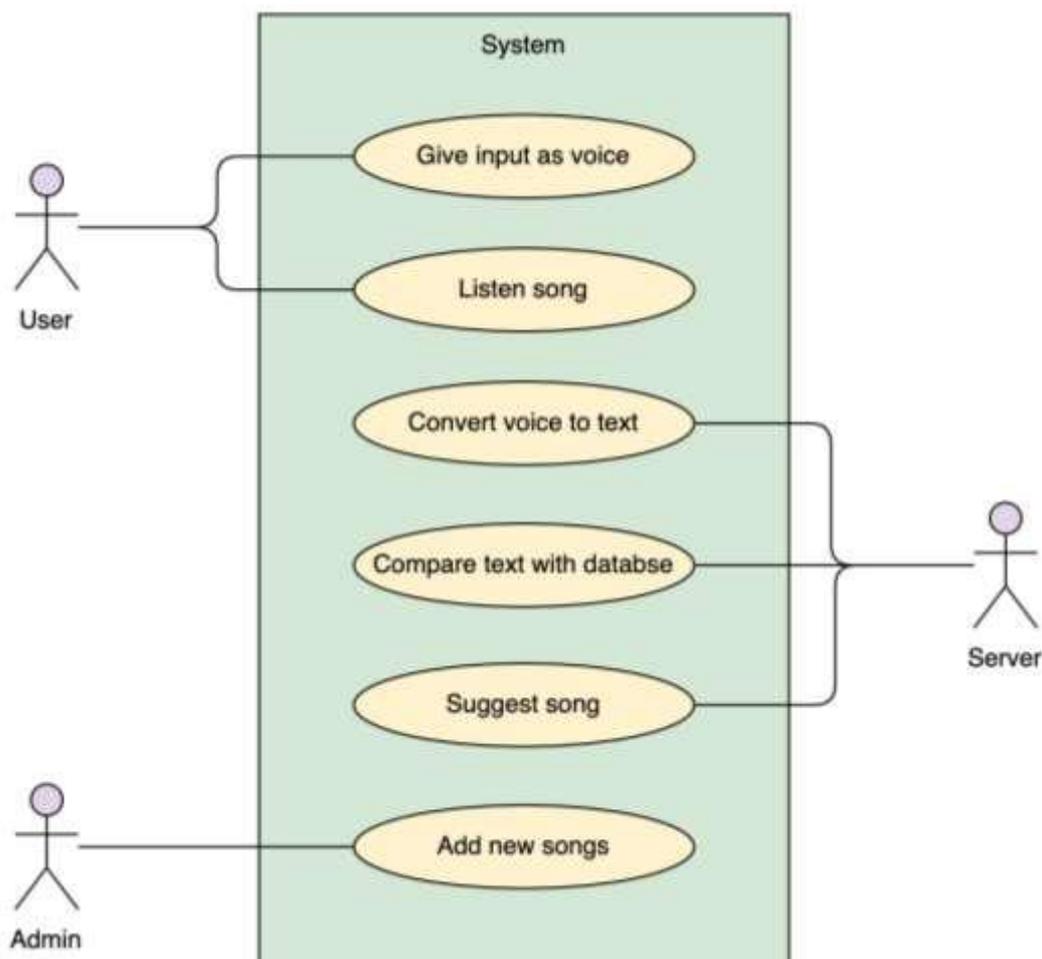
2. LITERATURE REVIEW:

A literature survey is a text of a scholarly paper that contains current information, including analytical observations, as well as theoretical and methodological contributions to a particular topic. Literature studies use secondary sources and do not mention any new or original experimental work. Supporting the Vector Machine algorithm. SVM can be used for classification or regression problems. It considers an optimal boundary between potential outputs. The training dataset that we used is the Olivetti face that includes 400 faces and its ideal values or parameters.. This paper discusses the probability-based music mood model and the implementation of the music recommendation method using the mood model. We're introducing three types of mood-based music recommendation services, for PC and mobile devices, and the internet. We take the customer satisfaction check for three months via a web-based music recommendation service.

3. PROJECT DESIGN:

Use Case Diagram of Emotion Detection Using Voice & Song Suggestion:

Following figure shows the use case diagram of Voice Based Music System. There are actors - Admin, Server and User and they perform the respective task as shown in the oval.

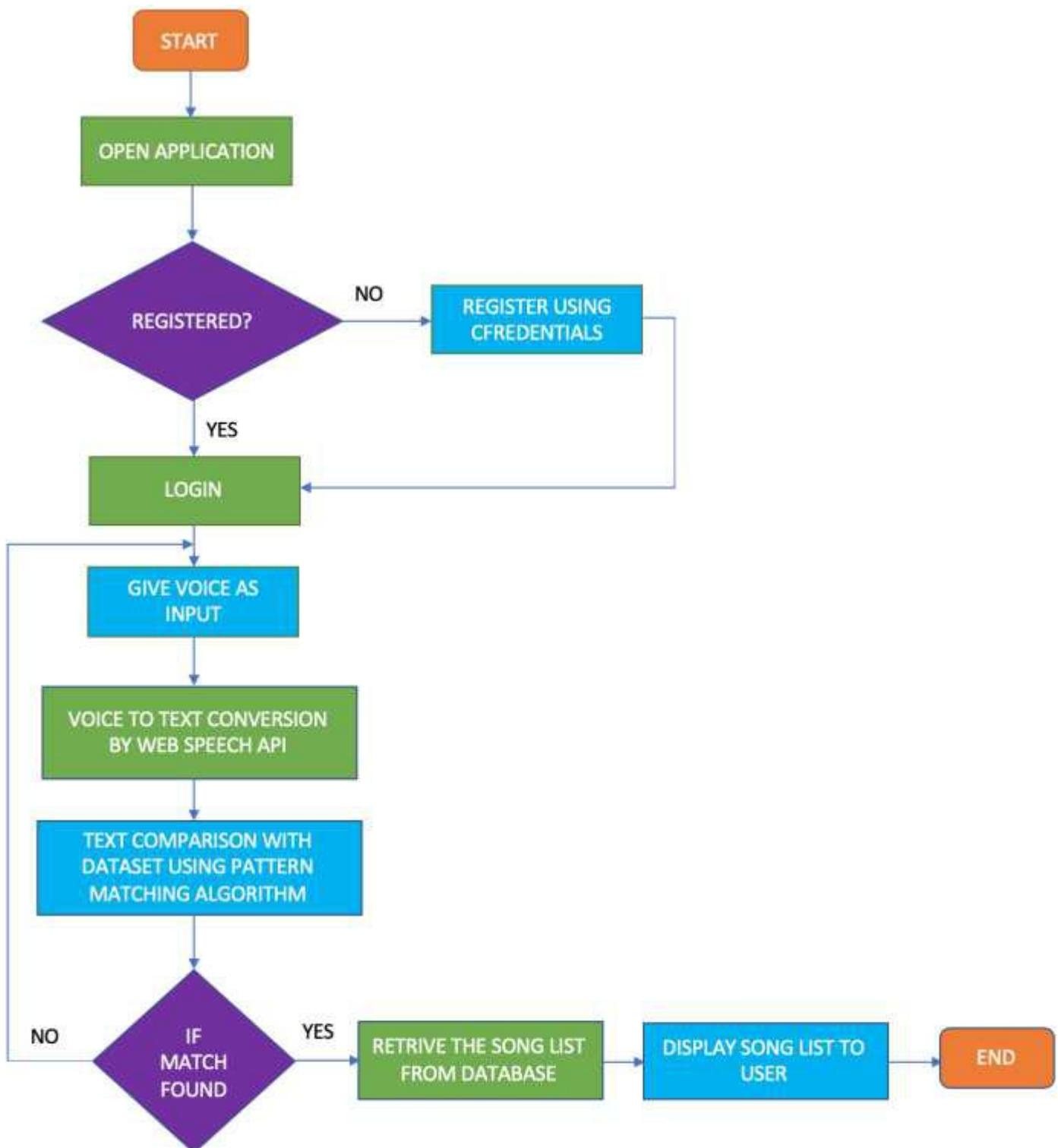


- **User:** Open the web application . Provide voice as input to system.
- **Admin:** Manipulate and managing of the system database by adding songs ,removing songs etc.
- **System:** Takes the voice as input from the user. It convert voice to text and compares the interpreted text with words in system. If finds the match with text then display the songs list related to users mood.

3.1. Flowchart:

First the user will open the web application. The user has to login using the valid credentials like username and password. If user is new to the application, he can register with us and then login. Once logged in user can give input as voice to the system. The system will convert this voice into text using the Web Speech API. The user will search the songs according to their mood and system will give output as mood based songs. The user has also given an option to like the song videos and user can see the list of liked videos in their profile. This application simply enhance the mood of user by suggesting the mood based songs. Speech recognition requires processing speech via the microphone of the user, which is then tested by speech recognition services.

- **Start() Function:** When the start process is called, it reflects the moment the web application needs to start recognition. When the speech data streams live via the data media source, Then this start call reflects a moment in time where the service must begin to listen and try to suit the grammar associated with this order. If the program has successfully listened to the acknowledgement, the user agent will trigger the case.
- **Stop() Function:** The stop approach is an order for the recognition service to avoid listening to more audio, and to try to return the result using only the audio that has already been provided for this recognition.
- **Abort() Function:** The abort approach is an instruction to stop listening immediately and to stop remembering and not to return any information but to do so. If the abort form is named, the speech service must avoid understanding it.



4. WEB SPEECH API INTERFACES:

The Speech Synthesis Interface of the Web Speech API is the speech service interface controller, which can be used to retrieve the synthesis voice information available on the computer, start and pause speech, and other commands besides.

- `SpeechSynthesis.cancel()`: Eliminates all utterances from the list of utterances.
- `SpeechSynthesis.getVoices()`: returns a list of `SpeechSynthesisVoice` objects that represent all available voices on the current computer.
- `SpeechSynthesis.pause()`: places the `SpeechSynthesis` object in a paused state.

- `SpeechSynthesis.resume()`: returns a list of `SpeechSynthesisVoice` objects that represents all available voices on the current device.
- `SpeechSynthesis.speak()`: adds a phrase to the list of utterances; it will be spoken when each other's utterance is queued before it is spoken. Axios is a promise-based HTTP client that works both in the browser and node.js environment.

4.1. AXIOS NPM:

Axios is a promise-based HTTP client that works in both the browser and the node.js environment. It provides a single API for dealing with XMLHttpRequests and the HTTP interface of the server. Also, it wraps up requests using the ES6 new commitment syntax polyfill. We need to install it first before using Axios.

- `$ npm install axios`

4.2. FEATURES OF AXIOS NPM:

- Make XMLHttpRequests from the browser
- Make http requests from node.js
- Supports the Promise API
- Intercept request and response
- Transform request and response data
- Cancel requests
- Automatic transforms for JSON data
- Client side support for protecting against XSRF

5. RESULTS AND DISCUSSIONS:

The web application SongPro was developed using Hypertext Markup Language (HTML), Cascading Style Sheet (CSS), JavaScript Object Notation (JSON), Structure Query Language (SQL), Bootstrap Theme, JavaScript.

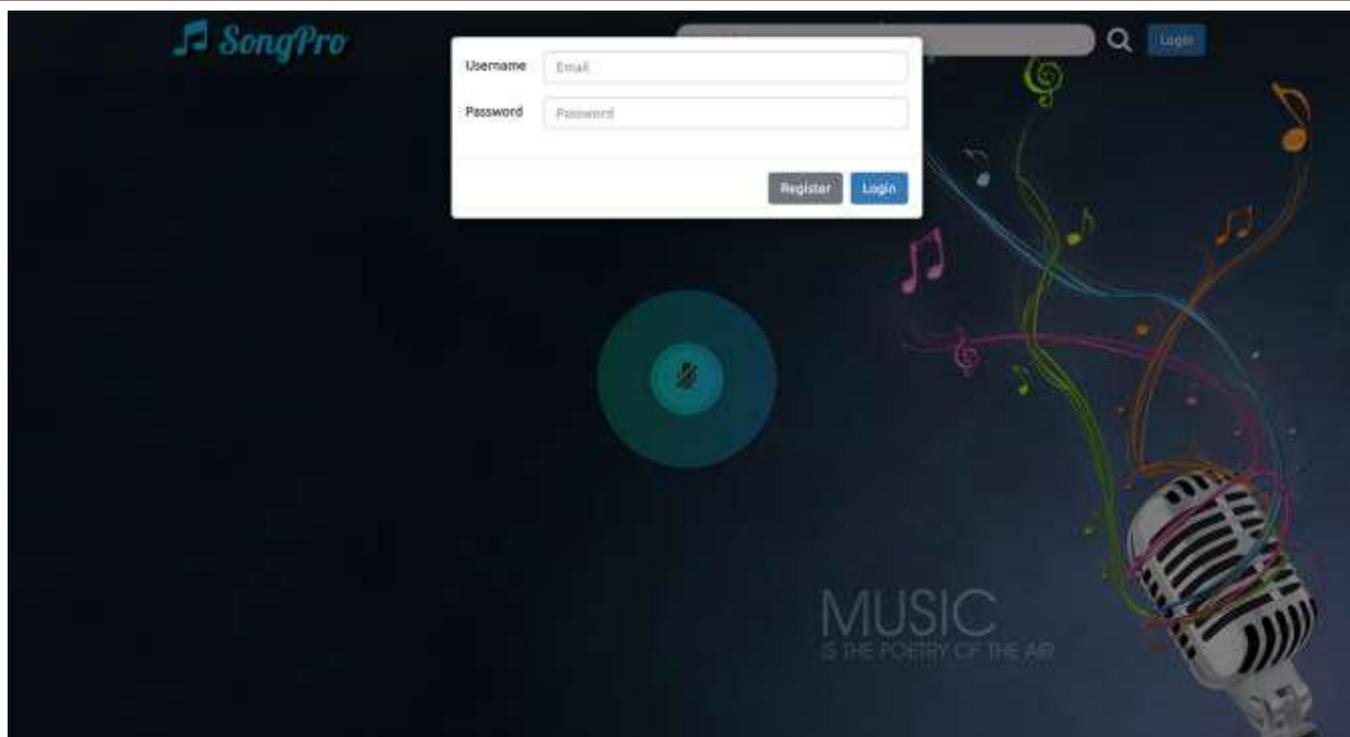
5.1. Application Home Page:

The web application home page displays the SongPro webpage. The page has the header which has options such as Search for songs, Login or Register.



5.2. Register and Login Page:

After clicking on Login button, It displays the login page for current users at the top of the application and a path for new users to register quickly onto the application.



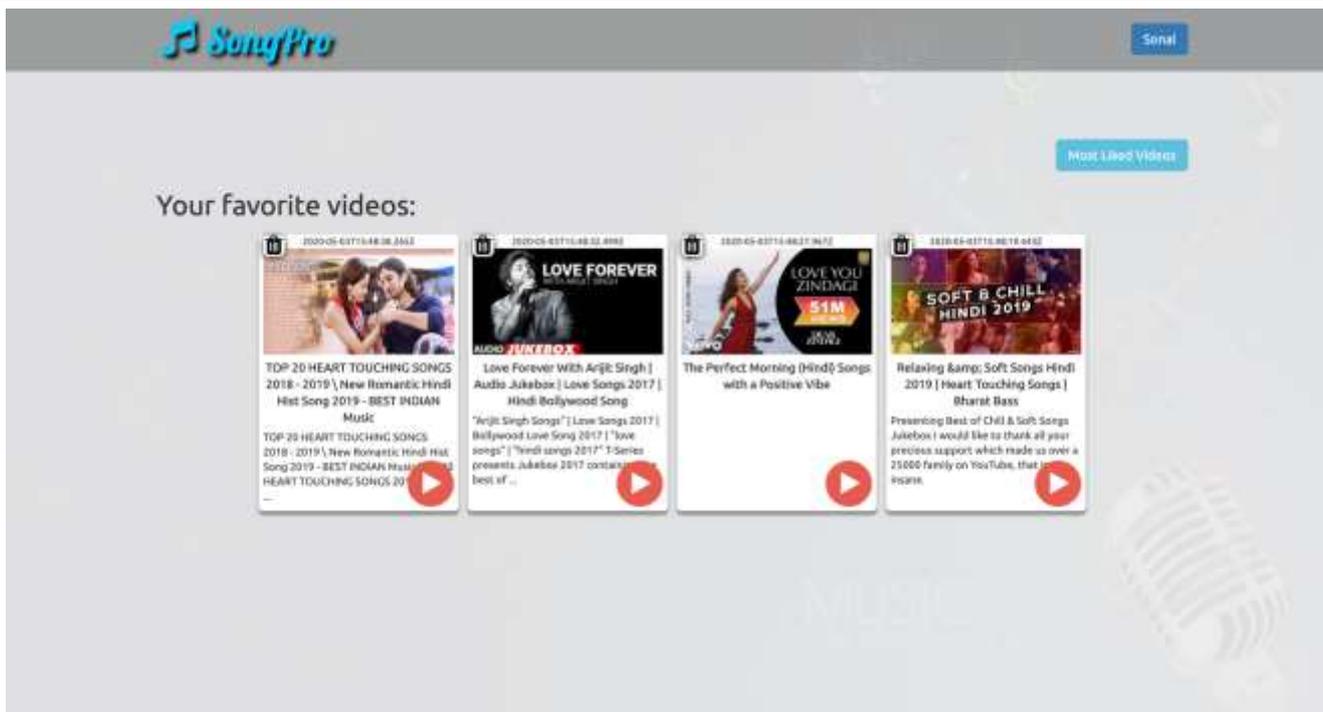
5.3. User Home Page (after the user has signed in):

The user home page has the header which has options such as Search for song, Your Favourite, Most Liked Videos. User name has log out option. User can give input as a voice to the system. The system will convert this voice into text using the web speech API and show it on the search bar. After that when the user clicks the search button it will show the list of songs related to his/her voice input.



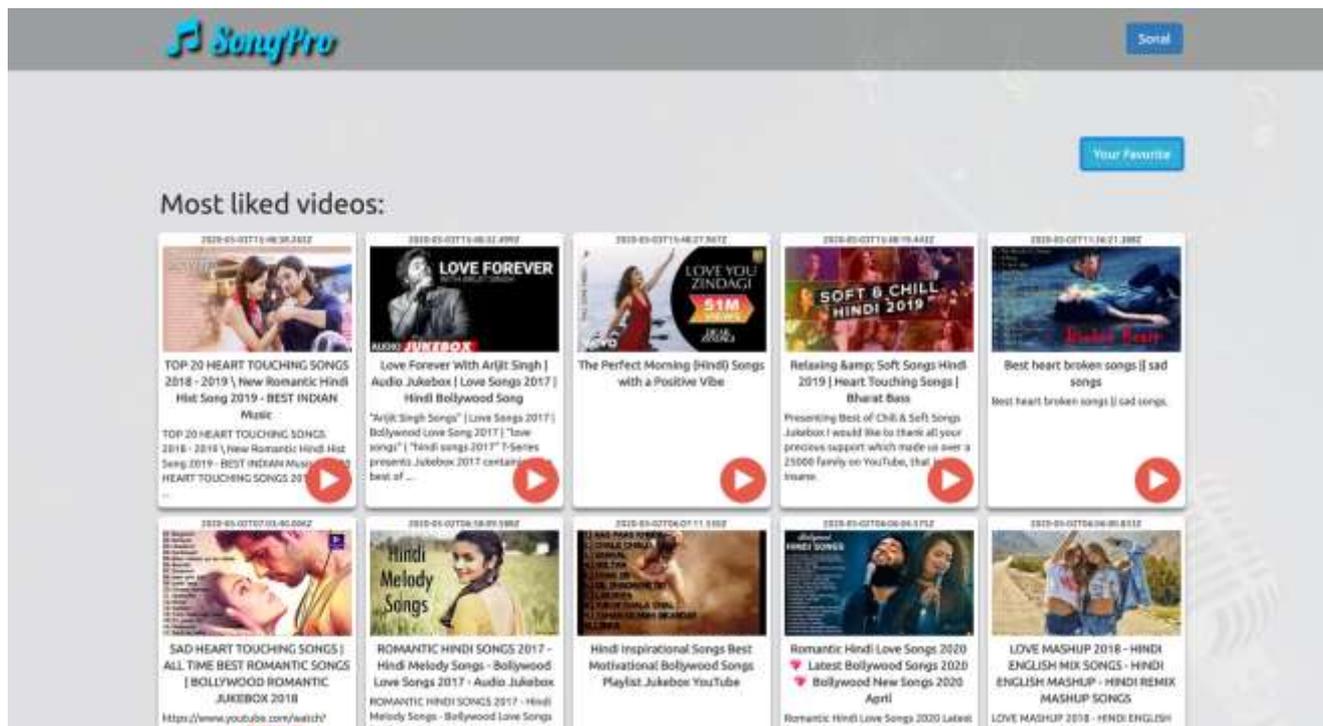
5.4. Users Favourite Videos :

The web application Favourite videos page displays the favourite videos of that particular user. It also displays the logout option at top of the application and a path to see most liked videos of by the user.



5.5. Most Liked Videos:

The web application Most Liked Videos page displays the most liked songs by the user. It also displays the logout option at top of the application and a path to see the user’s favourite songs/videos.



6. CONCLUSION:

The Emotion Detection and Song Suggestion program is used to simplify and give the end-user a better music player experience. The program meets the basic needs of music listeners without disrupting them as current applications do. It uses technology to improve the system's contact with the user in several respects. This uses a voice to a text converter to track the user's emotions. The user will also be informed of emotion-based songs that are not being played to help them free up storage space.

7. ACKNOWLEDGMENT: We wish to express our heartfelt appreciation and deep sense of gratitude to our project guide **Prof. J. A. Gaikwad** for her encouragement, invaluable support, timely help, lucid suggestions and excellent

guidance which helped us to understand and achieve the project goal. Her concrete directions and critical views have greatly helped us in the successful completion of this work. We are also thankful to all those who helped us directly or indirectly in the completion of this work.

REFERENCES:

1. Hemanth P, Adarsh , Aswani C.B, Ajith P, Veena A Kumar “EMO Player”, Students, Assistant Proffesor, Department of Computer Science,Saintgits College Of Engineering, Pathamuttom, Kottayam, Kerala, India Issue: 04 | Apr-2018
2. JungHyun Kim, Seungjae Lee, WonYoung Yoo, "Implementation and analysis of mood-based music recommendation system", 2013 15th International Conference on Advanced Communications Technology (ICACT), pp. 740-743, 2013.
3. M. K. Sarker, K. M. R. Alam, M. Arifuzzaman, "Emotion recognition from speech based on relevant feature and majority voting", 2014 International Conference on Informatics Electronics Vision (ICIEV), pp. 15, 2014.
4. <https://developers.google.com/web/updates/2013/01/Voice-Driven-WebApps-Introduction-to-the-Web-Speech-API>
5. <https://youtu.be/sHeJgKBaiAI>
6. <https://youtu.be/4npEhlacoHU>
7. <https://youtu.be/sHeJgKBaiAI>