

## A Survey on Speech Emotion Based Music Recommendation System

Chirag Kothawade<sup>1</sup>, Gourie Jagtap<sup>2</sup>, PreetKaur Relusinghani<sup>3</sup>, Vedang Chavan<sup>4</sup>, Prof.Smitha S Bhosale<sup>5</sup>

<sup>1234</sup> Student, <sup>5</sup> Assistant Professor,

Department of Computer Engineering, Sinhgad Institute of Technology and Science, Pune,  
India

**Abstract** - Psychological research has proven that music relieves stress, elevates mood, and is responsible for the release of “feel-good” chemicals like oxytocin, serotonin, and dopamine. It comes as no surprise that music has been a popular tool in rehabilitation centers and therapy for various disorders, thus with the interminably rising numbers of people facing mental health-related issues across the globe, addressing mental health concerns is more crucial than ever. Despite the existing music recommendation systems, there is a dearth of holistically curated algorithms that take care of the needs of users. Given that, an undeniable majority of people turn to music on a regular basis and that music has been proven to increase cognition, memory, and sleep quality while reducing anxiety, pain, and blood pressure, it is the need of the hour to fashion a product that extracts all the benefits of music in the most extensive and deployable method possible. Our project aims to ameliorate our users’ mental state by building a comprehensive mood-based music recommendation system called “Viby”.

**Keywords** Language, Communication, Speech recognition, Interaction.

### I. INTRODUCTION

Music has been around since the very beginning of civilizations and languages themselves. Humans have turned to music in times of grief, celebration and every event that lies in between, be it mundane or poignant. It not only affects our mental state of mind but has scientifically been shown to have positive physical effects on our bodies. Despite users being spoilt for choice when it comes to

music apps, there is a shortfall of software that seeks to improve the user’s mental state by recognizing their emotions. The intersection of music and music recommendation systems that strive to provide

users with personalized and enjoyable listening experiences. Traditionally, these systems [7] have relied on user listening history, preferences, and song metadata to make recommendations. However, they often overlook a crucial aspect of music appreciation—emotions. Music has a profound impact on our emotions and can be a powerful tool for expressing and managing them. People often listen to music that aligns with their current emotional state or helps them shift to a desired emotional state.

#### A. Motivation

The Goal is to enhance the user’s music listening experience by accurately detecting their emotional state from speech cues and recommending music tracks that align with those emotions.

#### B. Problem Statement

Speech Emotion-based Music Recommendation System revolves around developing a technology that accurately detects and interprets the emotional content in a user’s speech and utilizes this information to suggest music tracks that match their emotional state.

#### C. Relevance

The project is intended for people who are passionate about music and are constantly looking for new tracks, genres, or artists to explore. video.

#### D. Promote Efficiency

Efficiency is crucial for the effective

functioning. We aimed to reduce manual workloads and free up time for listeners.

#### E. Objectives

- A. To enhance user satisfaction, engagement, and the overall listening experience
- B. To leverage the emotional cues present in a user's speech.
- C. To deliver personalized music recommendations that resonate with their current emotional state.
- II. To explore new technologies and approaches to enhance the system's effectiveness [4].

#### LITERATURE REVIEW

The goal of a "Speech Emotion-based Music Recommendation System" is to enhance the user's music listening experience by accurately detecting their emotional state from speech cues and recommending music tracks that align with those emotions. Performing a speech emotion based music recommendation system involves the integration of speech emotion analysis with music recommendation algorithms.:

#### A. Acoustic Features and Classification for speech emotion based music recommendation system

In the paper [1], A. Kanjirath, N. Hossain, S. Madria explored the connection between speech emotion recognition and affective music classification, emphasizing how speech emotions can influence music recommendation. Its technique was to analyze the acoustic features of speech to detect emotions and recommend music based on these emotions.

#### B. Lyric Analysis for speech emotion based music recommendation system

In the paper [3], the author focused on integrating emotions in music recommendation systems, including speech emotion analysis, and explores lyrics as a source of emotional cues. Its technique was to analyze the lyrics of songs and user-generated content to understand and recommend music that matches emotional content

#### C. User Reviews and Sentiment Analysis for speech emotion based music recommendation system

In User Reviews and Sentiment Analysis, [2], the author investigates various sources of emotional cues, including user reviews, to recommend music based on emotions. Its technique was to analyze user reviews and sentiment to understand emotional associations with music and make recommendations accordingly

#### D. Hybrid Recommendation Systems for speech emotion based music recommendation system

For Hybrid Recommendation Systems, [4] presented a hybrid recommendation system that combines emotion and audio content analysis for improved recommendations. Its technique was to combine multiple data sources, such as speech emotion analysis, audio features, lyrics, and user behavior, to create a more robust recommendation system.

#### III. METHODOLOGY

Our project is designed as a web-based, comprising a client-side interface and a server-side backend. The system architecture includes components for listeners management, deaf management and enthusiastic management, and user authentication.

##### Data Collection

Data is collected from users via user-friendly interfaces, with secure storage in a database for processing.

This streamlined approach ensures accurate and efficient data management.

#### A. Feature Extraction

Mel-frequency cepstral coefficients (MFCCs): Represent the short-term power spectrum of a sound.

Prosodic features: Include pitch, intensity, duration, and rate of speech. Formant frequencies: Key frequencies in the speech signal related to vowel production.

#### B. Emotion Models

Gaussian Mixture Models (GMMs): For modeling the distribution of features for each emotion.

Support Vector Machines (SVMs): Used for classification based on extracted features.

Deep Learning models: Such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs).

#### C. Technologies Used

We have utilized modern web technologies such as HTML5, CSS3, and JavaScript to develop user interfaces, ensuring cross-browser compatibility and responsive design (as suggested in [6]).

To ensure data security, we have implemented

the JWT technique for user authentication to check the role of the user and encryption techniques for sensitive data storage.

#### IV. PROJECT DETAILS

1) *Introduction:* A Speech Emotion-Based Music

Recommendation system combines with Speech emotion recognition with music recommended algorithms.

2) *Scope:* Speech Recognition, User Authentication and Profiles, Music Database Integration, User Interface, Facial Expression Analysis, Music database Recommendation Algorithm.

Developing a Technology that accurately detects emotional content in users' speech and suggest music.

3) *Software Requirements:*

Operating system: Windows.

Coding Language: Python.

IDE: Spyder.

Database: SQLite3.

4) *Hardware Requirements:*

System: Intel I5 Processor

Hard Disk: 256 GB

Monitor: 15

RAM: 8 G

5) *Evaluation Metrics:* Evaluation metrics play a crucial role in assessing the performance of a speech emotion recognition (SER) system like Accuracy, Precision, Recall, Confusion Metrics.

6) *System Architecture:* The project's goal is to recognize and hear a person's voice in order to recognize their words. The system can tell if a person's voice is unhappy, pleased, or angry by listening to it. The necessary information regarding who is expected to have those needs, as well as what capabilities the voice interface should have to suit those demands.

7) *Limitations:* Variability in Speech Emotion Recognition, the accuracy of speech emotion recognition can be affected by individual differences, accents, and noisy environments, posing a challenge to precise emotional analysis.

8) *Challenges:* Data Privacy Concerns, Collecting and analyzing speech data for emotion can raise privacy concerns, necessitating robust data protection measures and user consent.

9) *Applications:* Music streaming services, Health and Wellness, Entertainment and Gaming, Voice Assistants, Education and E-learning etc.

10) *Results:* Created a system which will suggest or recommend songs on the basis of speech and facial expressions.

11) *User Classes and Characteristics:* The end users of a music recommendation system based on speech can be diverse and can encompass a wide range of individuals who enjoy music and prefer voice interaction for discovering new songs or playlists.

12) *System Implementation Plan:* All the functional and non-functional requirements along with system requirements are stated above. Iterative model will be followed for the system build and the modules will be accordingly built, tested, and integrated will the system. After the whole implementation is done overall system will be tested again.

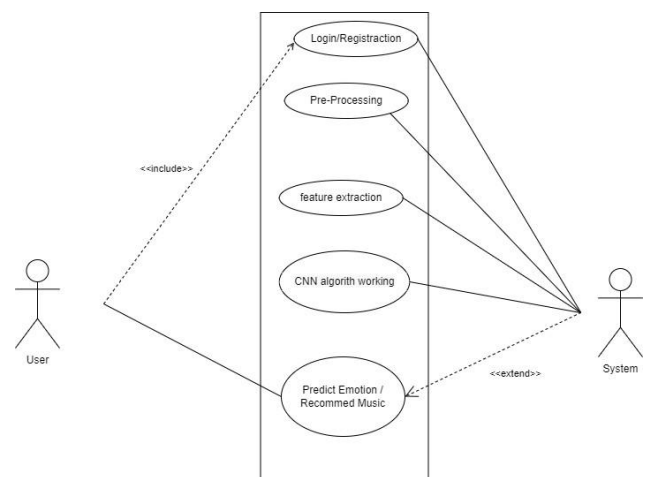


Figure 1: Block Diagram

#### V. DISCUSSION

The Role of Emotions in Music: Emotions are an integral part of the human experience, and music has a unique ability to resonate with and evoke specific emotional states. Whether it's the joy of an upbeat pop song, the nostalgia of

a classic ballad, or the tranquility of ambient sounds, music has the power to enhance or mirror our emotions. Speech Emotion-Based Music Recommendation System: The system [1] analyzes the user's speech patterns, tone, and emotional cues in real time. By discerning the user's emotional state from their voice, the system can recommend music that resonates with their feelings at that moment. Also, system analyzes users' emotions based on facial expressions and recommends music based on the user's mood.

#### *A. Future Directions*

As the project advances, potential enhancements include integrating SER with LMS platforms for synchronized data exchange [5], implementing AI algorithms for personalized learning, developing mobile applications for broader accessibility, exploring blockchain for enhanced data security, and adapting SER for global use with multilingual support. Additionally, adapting SER to accommodate multiple languages will broaden its global utility and accessibility.

Table 1: A SUMMARY OF RESEARCH REVIEWED

Author /Year of Publication	Title	Strength	Weakness
Taiba Majid Wani, Teddy Surya Gunawan, Syed Asif Ahmad Qadri, Mira Kartiwi, Eliathamby Ambikairajah, 2021	A Comprehensive Review of Speech Emotion Recognition Systems	Integrates facial emotion analysis into music recommendation.	Limited coverage of real-world applications.
Ashwin V. Gatty, G. S. Shivakumar, Kiran Shetty, 2021	Speech Emotion Recognition using Machine Learning	Provides a comprehensive overview of affect recognition, which includes facial expressions for music recommendation.	Lacks specific focus on music recommendation.
Ziyang Yu, Mengda Zhao, Yilin Wu, Peizhou Liu, Hexu Chen, 2020	Research on Automatic Music Recommendation Algorithm Based on Facial Micro-expression Recognition	Incorporates EEG signals for real-time music recommendation based on emotions.	Limited discussion on the scalability and user adoption.
Huihui Yang, Yi Liu, and Ying Tang, 2019	Emotion-Based Music Recommendation: Audio, Lyrics, and User-Review	Offers insights into music emotion recognition techniques that can be incorporated into recommendation systems.	Emphasis on emotion recognition more than the direct recommendation.
Subhradeep Kayal, Shiva Sundaram, and Anirban Dutta, 2018	Speech Emotion Recognition for Affective Music Classification: An Experimental Study	Provides an in-depth analysis of music emotion recognition methods, which can serve as a foundation for emotion-based recommendation.	Doesn't delve deeply into the technical aspects of recommendation.

Saikat Basu, Jaybrata Chakraborty, Arnab Bag and Md. Aftabuddin, 2017	A Review on Emotion Recognition using Speech	Offers an understanding of the connection between music and emotion, which is fundamental for emotion-based music recommendation.	Lacks specific recommendations on system implementation.
Alexandros Nanopoulos, Anastasios Tefas, and Ioannis Karydis, 2017	Towards an Automatic Classification of Music Emotional Content: Unsupervised Clustering of a Million Songs	Discusses individual differences in music perception and emotion, which can inform personalized recommendation systems.	Limited focus on technical aspects of recommendation.



## VI. CONCLUSION

In summary, the Speech Emotion-Based Music Recommendation System revolutionizes music recommendations by understanding and adapting to users' emotional states through speech analysis. It enhances user experience, engagement, and personalization while maintaining privacy and ethics. As it continues to evolve, it holds the potential to deeply connect users with music on an emotional level, shaping the future of music technology.

## ACKNOWLEDGEMENTS

We are deeply grateful to Prof. Smitha S Bhosale sir, our respected Head of the Department (HOD), and our department for the support, guidance, and help provided during the research conducted for the survey paper. Their expertise, encouragement, and mentorship have been essential in shaping our journey. Their unwavering belief in our capabilities and commitment to our academic growth has been a source of inspiration. We truly appreciate their dedication to our success. Thank you for being pillars of knowledge and support, and for your pivotal roles in our project's progress.

## REFERENCES

- [1] Taiba Majid Wani, Teddy Surya Gunawan, Syed Asif Ahmad Qadri, Mira Kartiwi, Eliathamby Ambikairajah. "A Comprehensive Review of Speech Emotion Recognition Systems" In IEEE, 2021
- [2] Ashwin V. Gatty, G. S. Shivakumar, Kiran Shetty. "Speech Emotion Recognition using Machine Learning" In IJRESM, 2021
- [3] Ziyang Yu, Mengda Zhao, Yilin Wu, Peizhou Liu, Hexu Chen. "Research on Automatic Music Music Recommendation Algorithm Based on Facial Micro-expression Recognition" In CCC, 2020.
- [4] Huihui Yang, Yi Liu, and Ying Tang. "Emotion-Based Music Recommendation: Audio, Lyrics, and User-Review" In 2019.
- [5] Subhradeep Kayal, Shiva Sundaram, and Anirban Dutta. "Speech Emotion Recognition for Affective Music Classification: An Experimental Study" In 2018.
- [6] Saikat Basu, Jaybrata Chakraborty, Arnab Bag and Md. Aftabuddin. "A Review on Emotion Recognition using Speech" In IEEE, 2017.
- [7] Alexandros Nanopoulos, Anastasios Tefas, and Ioannis Karydis. "Towards an Automatic Classification of Music Emotional Content: Unsupervised Clustering of a Million Songs" In 2017.