JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EXPRESSIFY: EMOTION-BASED MUSIC RECOMMENDATION SYSTEM

[1]Pooja Hansda, [2]Priya M, [3]Radhika D, [4]Sharadhi S Shetty

[1234]Final Year B.E. Students, Department of Computer Science & Engineering,
Srinivas Institute of Technology, Mangaluru, India

Abstract: Across cultures, music plays a vital role in reflecting and shaping human emotions, often transcending language and personal barriers. It has the ability to uplift, comfort, energize and soothe and sometimes all at once. However, in today's world of music recommendation systems, we are left with suggestions that feel impersonal or disconnected from our current emotional state. This is where Expressify comes in. Expressify is an emotion- based music recommendation system designed to offer a deeply personalized listening experience. By analyzing real-time emotional cues - such as facial expressions, speech tone, and text sentiment. Expressify matches the mood with the perfect music genre and track. Using a curated dataset with emotion-genre associations. Expressify delivers accurate, mood-aligned recommendations, whether you're unwinding or seeking motivation. This innovative approach not only makes music discovery more dynamic but also more meaningful, offering a truly emotional and personalized experience for every listener, highlighting the potential of integrating emotion-aware AI with entertainment platforms to create more engaging and meaningful user interactions.

Index Terms **- Deepface, CNN, VaderSentiment, TextBlob, Parselmouth.**

I. INTRODUCTION

Music serves as a strong channel for conveying emotions and shaping how people feel. With advancements in technology, especially in artificial intelligence and human-computer interaction, creating personalized experiences has become a significant focus area. EXPRESSIFY, an emotion-driven music recommendation system, is developed to connect a user's current emotional state with suitable music choices, providing a more personalized and mood-based listening experience. The proposed system detects user emotions dynamically by processing facial gestures captured through a live camera feed, using pre-trained emotion recognition models. These features operate independently, allowing the system to recommend music either by detecting facial emotions or analyzing vocal tones. By integrating these capabilities with user interaction data and mood analysis, the system generates highly accurate music recommendations tailored to the user's current emotional state. This intelligent recommender system leverages cutting-edge technologies such as machine learning, computer vision, sentiment analysis, and speech processing to create a seamless user experience. Its real-time capabilities, combined with its ability to interact and adapt based on user preferences, make it an innovative solution in the realm of music streaming applications.

The proposed solution introduces a real-time, mood-based music recommendation mechanism. Using emotion recognition through facial, speech, tone of text and mood analysis, the system evaluates the user's current emotional state through live video feed and brief interactive prompts. Key Features include, Facial Emotion Recognition (FER) which captures and analyzes facial expressions to present identify emotions like happiness, sadness, anger, or surprise etc. Speech and Text Emotion Recognition analyses the pitch of speech and polarity of text to detect emotions accurately. Mood-Based Music Selection to map detected emotions to corresponding genres, tempos, or artists, ensuring an emotionally resonant music choice. User Preference Learning continuously updates user profiles by analyzing listening history, preferences, and interaction patterns. Personalized Recommendations combines mood analysis with user preferences to provide recommendations that cater to both current emotions and long-term tastes. Real-Time Adaptability helps in providing instant adjustments to the playlist based on changing emotional cues.

Several previous studies have focused on analyzing emotions using individual methods such as voice input, facial imagery, or text sentiment separately. In contrast, our approach combines all three emotion-detection techniques into a single integrated system, utilizing multiple algorithms and libraries to ensure efficient and accurate functionality. The system emphasizes the use of lightweight, cost-effective models capable of recognizing the user's mood in real-time, categorizing newly added tracks, and recommending them accordingly [1]. One related system applies artificial intelligence and machine learning to provide personalized song suggestions by interpreting the emotional condition of the user. These systems often rely on features such as MFCC and deep learning techniques to connect vocal patterns with emotional states, achieving reliable performance across various datasets [2]. Another referenced study introduces an automated playlist generator that leverages facial expression data captured via a built-in camera, significantly reducing manual effort and system load while maintaining high accuracy. The proposed algorithm reportedly reaches 85–90% accuracy for real-time images and nearly 100% for static ones [3]. Machine learning algorithms including K-Nearest Neighbors (KNN), Logistic Regression, SVM, Random Forests and

Cosine Similarity, applied to Spotify dataset. The study aims to optimize accuracy and enhance user music preferences through algorithmic evaluation [4]. Using techniques like the MobileNet model with Keras, this system integrates seamlessly with Android for efficient emotion detection. The techniques used is Computer Vision and Machine Learning (ML) [5,6,8]. Convolution Neural Network used for emotion detection. This is managed using Pygame and Tkinter, reducing computation time and the overall cost while improving accuracy, the system is tested using the FER2013 dataset, where facial expressions such as happy, sad, angry, surprised and neutral are detected using an inbuilt camera [7]. The proposed system uses deep learning for emotion recognition and collaborative filtering for personalized suggestions [8]. Next model is using CNN and Haar Cascading algorithm which identifies user's emotions through facial expressions and offers music recommendations tailored to their mood and preferences. This method helps in reducing decision fatigue and enhances user experience by offering relevant music suggestions [9]. The proposed solution is similar to other emotion detection method using the facial expressions to analyze and based on that give the suggestions on music recommendations [10].

II. PROPOSED SOLUTION

Expressify identifies a user's current emotion through visual, vocal, and written inputs to curate a personalized music experience.

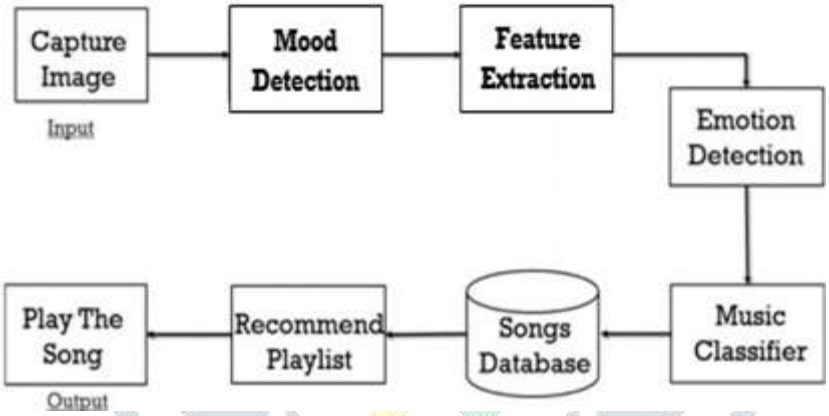

FIGURE 1: *The proposed system*

The system consists of the following major components. First is the user input interface which captures and processes the user's facial expressions and voice tones and emotion of text. Next comes the emotion detection module which analyses the captured input (either facial expression, voice tone and polarity of text) to detect emotions. Then the emotion genre mapping which maps the detected emotion to a relevant music genre. Recommendation engine suggests personalized music based on the emotion genre mapping. The music player delivers the selected tracks to the user for playback.

The Figure 2 displays how data moves through various components of the system. The process starts on the home screen, where the user is prompted to choose one of three input methods: opening the camera, recording their voice, or typing words. If the user selects the camera, the system performs face detection to analyze the mood. If the user chooses to record their voice, the system performs pitch analysis to determine the mood. If the user types words, the system conducts tone analysis for mood detection. After detecting the user's mood using any of the selected input methods, the system matches the emotion with relevant songs from its database and presents a playlist. The user then has the option to play the suggested tracks or initiate a new mood scan. If no rescan is selected, the system ends the session.

Tools and technologies used for emotion-based music recommendation system. Facial emotion detection is implemented using a deep learning framework designed to streamline expression analysis. Convolution Neural Network (CNN) used for emotion detection from images or videos due to its ability to capture spatial and hierarchical features effectively. Here the CNN uses 5 to 7 layers for feature extraction. SQLite and SQLAlchemy used for secure solution for managing user login details and dynamically tracking user emotions in an application. Text processing is handled through user-friendly natural language processing libraries, VaderSentiment designed for sentiment analysis offers polarity score ranging from -1 (negative) to 1 (positive) and googletrans facilitates language translation enabling the system to analyze text from various languages making it versatile. Acoustic features are extracted using Python-based tools that integrate with speech analysis platforms like Praat. Enabling seamless extraction and analysis of speech features such as pitch, intensity and voice quality.

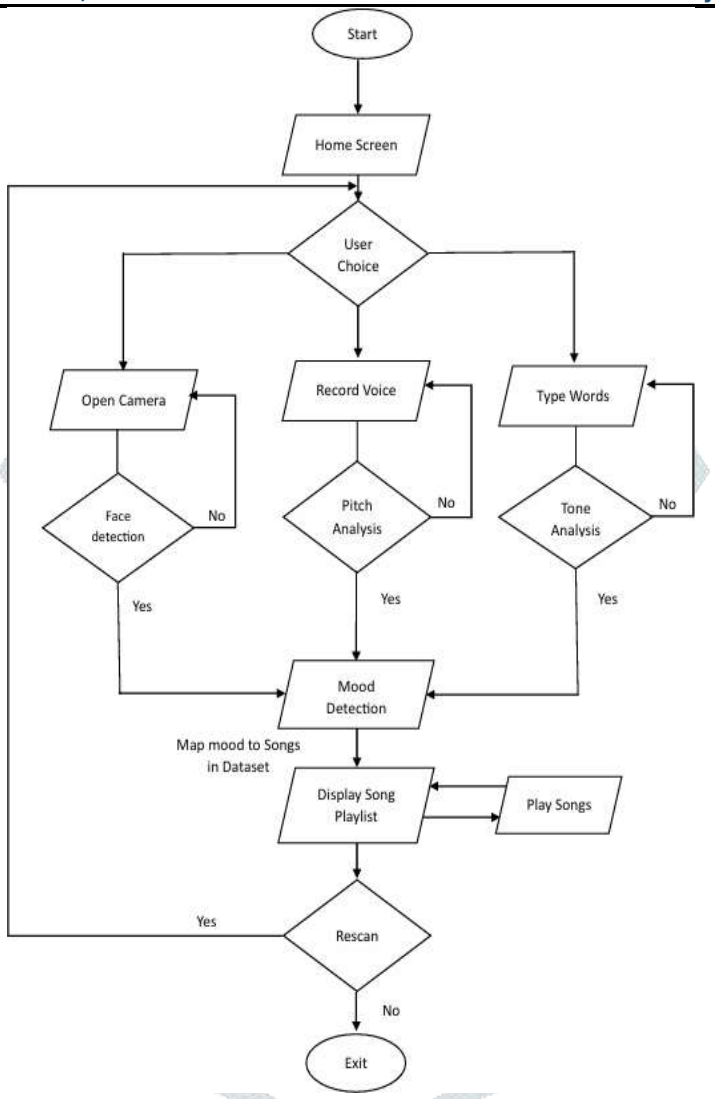

FIGURE 2: *System flow diagram*

III. RESULTS AND DISCUSSION

Emotion detection in the system is achieved by analyzing facial movements, vocal patterns, and the sentiment conveyed through user-provided text. Facial expressions, such as smiles, frowns, or raised eyebrows, reflect key emotions, including happiness, sadness, anger, and surprise. The system decodes these minor variations to accurately infer the user's mood. Similarly, the tone of voice, including pitch and rhythm, offers insights into emotions, with cheerful tones signaling happiness and trembling voices hinting at fear or anxiety. Sentiment analysis of textual input evaluates the choice of words and tone to recognize feelings such as excitement or sadness.

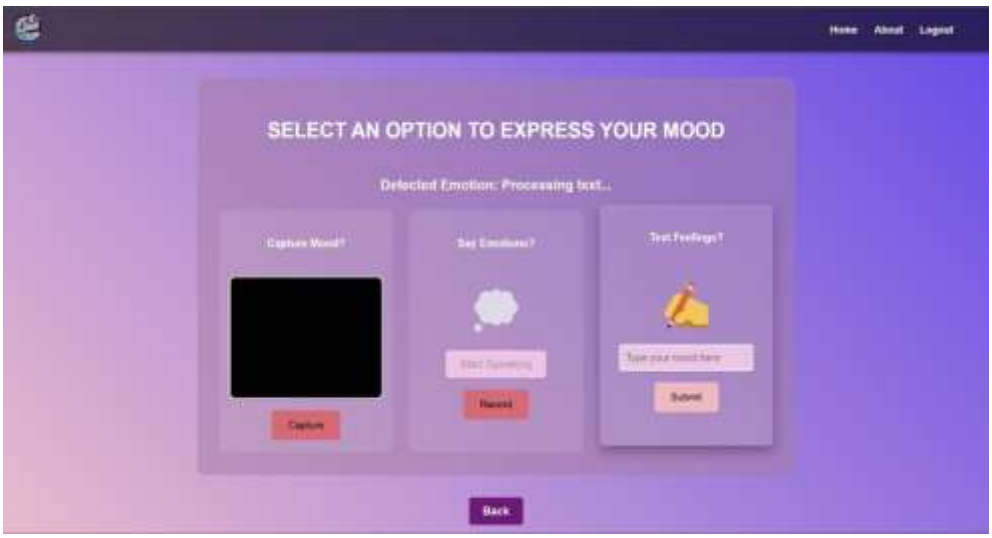

FIGURE 3: *Mood recognition page*

The figure 3 illustrates the mood recognition page which analyzes the input either image, voice or text to identify the emotions or sentiments expressed by the user. Here for example, we have given the text input as “I feel like I am in a happy mood” words like “happy” and phrases suggesting joy or contentment are detected. This process helps in understanding user’s emotional state based on their language.

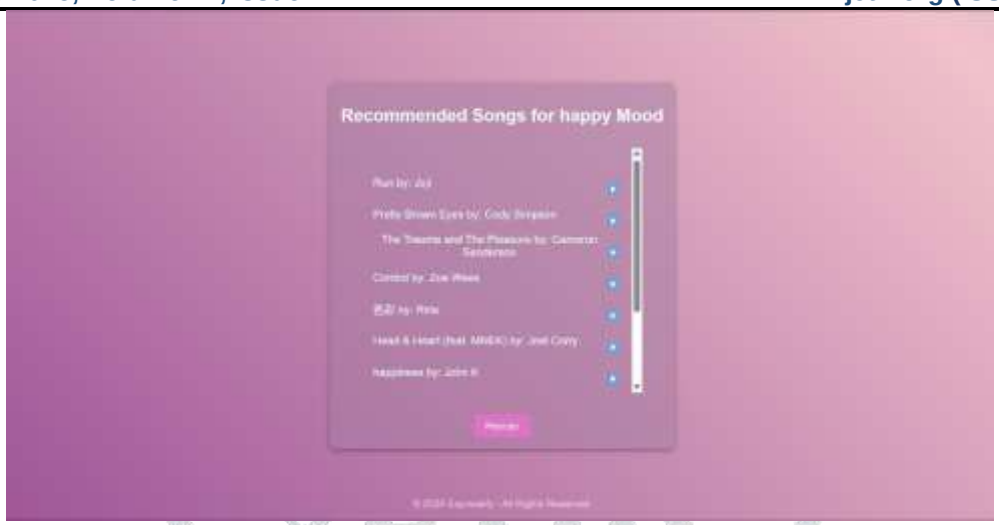

FIGURE 4: *Song suggestion page*

The figure 4 illustrates the detected emotion "happy" is identified in the third option by analyzing the tone of text, or the positive sentiment in text, reflecting feelings of joy and satisfaction and then recommending the curated music playlist for the detected emotion. The song when clicked redirects us to Spotify app where we can listen to the music as per our current mood. If the user presses the rescan button, then it is redirected to mood recognition page.

IV. Conclusion And Future Work

EXPRESSIFY showcases the powerful potential of merging facial emotion recognition and voice tone analysis to deliver music recommendations tailored to users' emotional states. By leveraging individual inputs—such as facial cues, vocal tones, and textual sentiment—the system ensures a smooth and engaging user interaction. Through the integration of cutting-edge technologies like machine learning, sentiment analysis, and computer vision, EXPRESSIFY offers a unique and responsive experience that connects emotions to music choices more intuitively.

To enhance emotional accuracy, future iterations will incorporate biometric signals such as heart rate variability, enabling a richer understanding of user states. This advancement will enable a deeper and more accurate reading of users' emotions. Furthermore, by expanding to incorporate sign language recognition and other multimodal inputs, the system aims to become even more inclusive and emotionally aware, adapting more precisely to individual needs and preferences.

References

[1] Singh, H. R., Upadhyay, A., Sutar, M., & Aamer, N. (2022). Emotion Based Music Recommendation System. Undergraduate Students, Department of Computer Engineering, TCOE, Mumbai University, Maharashtra, India.

[2] Rajesh, A., Pavani, T., Sai Naveen, R. V., Manikanta, P., & Sandeep, P. (2024). Song Recommendation Based on Voice Tone Analysis. Department of Computer Science and Engineering, Chalapathi Institute of Technology, Guntur.

[3] Vinay, P., Prabhu, R. T., Satish Kumar, B., Jayanth, P., & Suneetha, A. (2021). Facial Expression Based Music Recommendation System. Department of Computer Science & Engineering, KKR & KSR Institute of Technology and Sciences.

[4] Parthasarathy, A. (2023). Music Recommendation Using Machine Learning Algorithms. Department of Computer Science and Engineering, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani, Chennai, India.

[5] Bhutada, S., Sadhvika, C., Abigna, G., & Reddy, P. S. (2020). Emotion-Based Music Recommendation System. Information Technology Department, Sreenidhi Institute of Science and Technology, Yamna pet, Hyderabad.

[6] Mahadik, A., Jagan, V. B., Milgir, S., & Kavathekar, V. (2021). Mood-Based Music Recommendation System. Department of Information Technology, Don Bosco Institute of Technology, Mumbai, India.

[7] Athavle, M., Mudale, D., Shrivastav, U., & Gupta, M. (2021). Music Recommendation Based on Face Emotion Recognition. Department of Computer Engineering, NHITM, Thane, India.

[8] Navale, V., Phand, S. V., Patil, P. K., Rangat, D. V., & Sathe, R. J. (2024). FeelFlow: Music Recommendation System Based on Facial Expression. Department of Computer Engineering, All India Shri Shivaji Memorial Society's College of Engineering, Pune, India.

[9] Madhu, A. K., Vallippokkil, S., Sreelakshmi, K. S., Sojan, S., & Arya, T. J. (2024). EMOSOUND: An Emotion-Based Music Recommendation System. Department of Computer Science and Engineering, College of Engineering, Kidangoor, Kottayam, Kerala.

[10] Benke, M., Ayare, M., Khairnar, A., & Sonawane, D. (2023). Music Recommendation Based on Emotion Detection Through Facial Expression. Department of Information Technology, RMDSSOE, Pune, India.