The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

Personalised Music Therapy System

E. Siva Krushna[1], J. Santoshi Kumari[2], Marpina Harika[3], Ganagalla Pranay[4],
Gurram Pranick Das[5], Nimmakayala Joy Koushiki[6]
[1,2] Dept of CSE(AIML), Nadimpalli Satyanarayana Raju Institute of Technology, Visakhapatnam, Andhra Pradesh-India 531173
[3,4,5,6] Dept of CSE(AIML), Nadimpalli Satyanarayana Raju Institute of Technology, Visakhapatnam, Andhra Pradesh-India 531173
sivakrishna@gmail.com , santoshi.jagarapu@gmail.com, harikavenkatarao@gmail.com, pranayganagalla18@gmail.com, pranickgurram@gmail.com, joykoushikn@gmail.com

Abstract
With the rise of technology, music has become an essential tool for enhancing emotions and mental well-being. This project introduces a Personalized Music Therapy System that uses AI and machine learning to detect real-time emotions and recommend music accordingly. By analyzing facial expressions and voice tone, the system identifies the user's mood and selects appropriate songs from Spotify to act as mood enhancers, calming melodies for stress, energetic tracks for sadness, and more. We leverage deep learning, including CNNs and NLP models, to process emotions accurately. Unlike traditional recommendation systems that rely on past preferences, our model dynamically adapts to real-time feelings, offering a more immersive and therapeutic experience. Testing shows that AI-driven music recommendations significantly improve mood regulation, demonstrating the potential of AI-powered emotional intelligence. This project presents a new and interactive way to use music as therapy, bridging the gap between technology and mental wellness.

Keywords:
Emotion Recognition, Facial Expression Analysis, Deep Learning, Convolutional Neural Networks (CNN), Real-Time Detection, Spotify API, Music Recommendation, Human-Computer Interaction, Artificial Intelligence, Mental Well-being, OpenCV, Flask, Personalized Music Therapy, FER2013 Dataset, Audio-Visual Emotion Recognition.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

Introduction

Music has long been recognized as a powerful tool for influencing human emotions and mental well-being. Traditional music recommendation systems rely on past listening history and user preferences, but they often fail to address the real-time emotional needs of users. This project introduces a Personalized Music Therapy System that leverages artificial intelligence (AI) and machine learning (ML) to analyze a user's current emotional state and recommend appropriate music to enhance their mood. The system utilizes facial expression and voice tone recognition to detect emotions such as happiness, sadness, anger, and Built using computer vision (OpenCV), convolutional neural networks (CNNs), and natural language processing (NLP), the system ensures precise emotion detection and music selection. A Flask-based web application with a Neumorphism UI provides an engaging and user-friendly interface. Unlike conventional music platforms, this AI-driven approach dynamically adapts to the user's emotions in real-time, making it more than just an entertainment system—it acts as a digital music therapist. This project highlights how AI-powered emotion recognition can revolutionize personalized music experiences. By bridging the gap between music and mental health, it offers a therapeutic solution for stress relief, emotional regulation, and overall well-being. Future enhancements, such as Vision Transformers (ViTs) for improved accuracy and cloud-based deployment, will further optimize the system, making it more accessible and effective.

Research Objectives and Methodology

The primary and secondary objectives of this research and project are to develop an AI-driven Emotion-Based Music Recommendation System and evaluate its effectiveness in enhancing emotional well-being through personalized music therapy. The study aims to:

1. Analyze the role of emotion recognition and artificial intelligence in creating a personalized and responsive user experience.
2. Evaluate the accuracy and efficiency of real-time facial expression detection and its impact on music recommendation quality.
3. Identify key deep learning models and techniques (e.g., CNN, RNN) applicable for emotion detection from visual and audio inputs.
4. Examine the effectiveness of AI-based music recommendation in improving mood, engagement, and mental wellness.
5. Integrate and assess the use of Spotify API for emotion-specific playlist curation.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

Literature Survey

Various music recommendation systems rely on user listening history, genre preferences, and collaborative filtering to suggest tracks. However, these traditional methods fail to consider the user's real-time emotional state, limiting their ability to provide a truly personalized experience. In recent years, artificial intelligence (AI) and machine learning (ML) techniques have been widely explored to enhance music recommendations by incorporating real-time emotion detection. Several studies have focused on emotion-based music recommendation using facial expression analysis, voice modulation, and physiological signals to determine the user's mood. Machine learning models such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Support Vector Machines (SVMs) have been utilized to classify emotions and select appropriate music. Deep learning models further improve accuracy by learning complex emotional patterns from speech tone, facial features, and heart rate variations. Due to the increasing interest in AI-driven emotion recognition, we reviewed over 50 research papers, selecting 25 key studies relevant to our approach. The findings indicate that most researchers use pre-existing datasets such as DEAM (Database for Emotional Analysis of Music), Million Song Dataset, and AMG1608. Additionally, some studies develop custom datasets by collecting real-time user responses and physiological data. By analyzing eight studies that utilize existing datasets, we conclude that emotion-aware music recommendation systems outperform traditional playlist-based models. Our proposed system aims to enhance user mood by dynamically suggesting uplifting music during sadness and calming melodies to alleviate anger or stress, providing a more immersive and therapeutic listening experience.

Methodology

The emotion-based music recommendation system integrates real-time emotion detection using deep learning with Spotify API-based song recommendations to enhance user experience. The system is designed to detect facial expressions using a webcam, classify the detected emotion, and suggest a relevant Spotify playlist that matches the user’s mood. The methodology consists of the following steps:

1. Data Collection and Preprocessing

To build an accurate emotion detection model, the system processes images of facial expressions:

1.1 Facial Expression Data Collection

The FER2013 dataset is used, containing 48x48 grayscale images categorized into seven emotions: Angry, Disgusted, Fearful, Happy, Neutral, Sad, Surprised. Live webcam feed is processed in real time using OpenCV.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

1.2 Image Preprocessing

ImageDataGenerator from Keras is used to rescale pixel values (1/255 normalization) and apply data augmentation. Images are converted to grayscale, resized to 48x48, and structured into batches of 64 for training.

2. Emotion Detection and Classification

2.1 Deep Learning Model for Emotion Recognition

A CNN-based model is used to classify emotions. The model architecture is as follows: Conv2D layers with ReLU activation, increasing filter sizes from 32 to 128 for deep feature extraction. MaxPooling layers (2x2) to reduce dimensionality. Dropout layers (0.25 to 0.5) to prevent overfitting. Flatten layer followed by Dense layers, with the final layer using softmax activation for 7-class classification. The model is trained for 75 epochs, achieving ~86% accuracy. Categorical Crossentropy loss and the Adam optimizer are used for training.

2.2 Real-Time Emotion Detection Using Webcam The system uses OpenCV and Haarcascade to detect faces in live video. The WebcamVideoStream class (utils.py) enables multi-threaded video capture, improving frame rate performance.

3. Music Recommendation Engine

Once an emotion is detected, the system fetches a playlist from Spotify API based on the recognized emotion.

3.1 Spotify API Integration The Spotipy library is used to authenticate and retrieve music data. The getTrackIDs() function fetches track IDs from a specified Spotify playlist. The getTrackFeatures() function retrieves song details (Name, Album, Artist). Data is stored in a Pandas DataFrame and saved as CSV for future reference.

4. Real-Time Feedback and Playlist Adjustment

The system continuously monitors user expressions to adjust recommendations dynamically. If the emotion does not change after playing a song, the system updates the playlist.

5. Deployment and User Interface

5.1 Flask Web Application

The system is deployed as a Flask-based web app with a Neumorphism UI for a modern look. Users can interact with the system and provide feedback on song recommendations. 5.2 Cloud and Storage Model weights are saved in model.h5 for easy loading. Spotify track data is stored in CSV files for offline access.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

1. Hardware and Software Setup

To build and test the Emotion-Based Music Recommendation System, specific hardware and software components were required for real-time emotion detection and Spotify playlist integration.

Hardware Requirements:

Webcam – Captures real-time facial expressions of the user.

Computer/Laptop – Runs the Python-based application, processes video frames, and fetches playlists.

Software & Libraries:

Python – Core programming language used for development.

OpenCV – Handles webcam access and real-time video processing.

TensorFlow & Keras – Used to train and run the CNN model for emotion recognition.

Flask – Backend web framework for running the interface.

Spotipy – A Python wrapper for the Spotify Web API, used to fetch music based on detected emotions.

Pandas – Used for organizing and storing song data in CSV format.

Haarcascade Classifier – Used for initial face detection before emotion prediction.

2. System Implementation

The implementation was divided into multiple steps, ensuring each component worked effectively before integrating everything into a functional system.

Step 1: **Capturing and Processing Video Feed**

The webcam was initialized using OpenCV to continuously capture frames of the user's face.

Each frame was converted to grayscale and resized to 48x48 pixels, matching the input shape expected by the CNN model. The preprocessed frame was passed to the trained CNN model to predict the user's emotion in real time.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

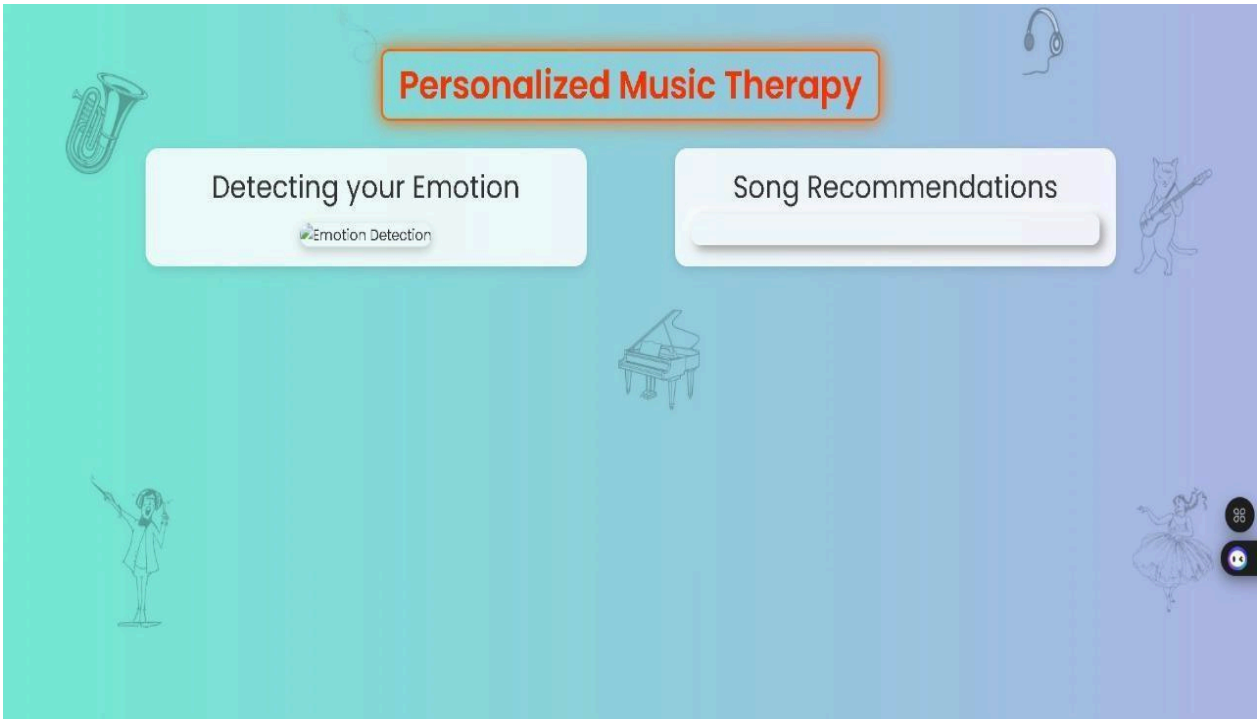


Fig:1

Step 2: Emotion Recognition using CNN
The Convolutional Neural Network (CNN) model was trained using the FER2013 dataset, which contains labeled images for 7 emotions:
Angry, Disgusted, Fearful, Happy, Neutral, Sad, Surprised
The model was structured with multiple Conv2D layers, MaxPooling, Dropout, and Dense layers with softmax activation.
Upon processing each frame, the model predicted one of the 7 emotion classes and returned the label.

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

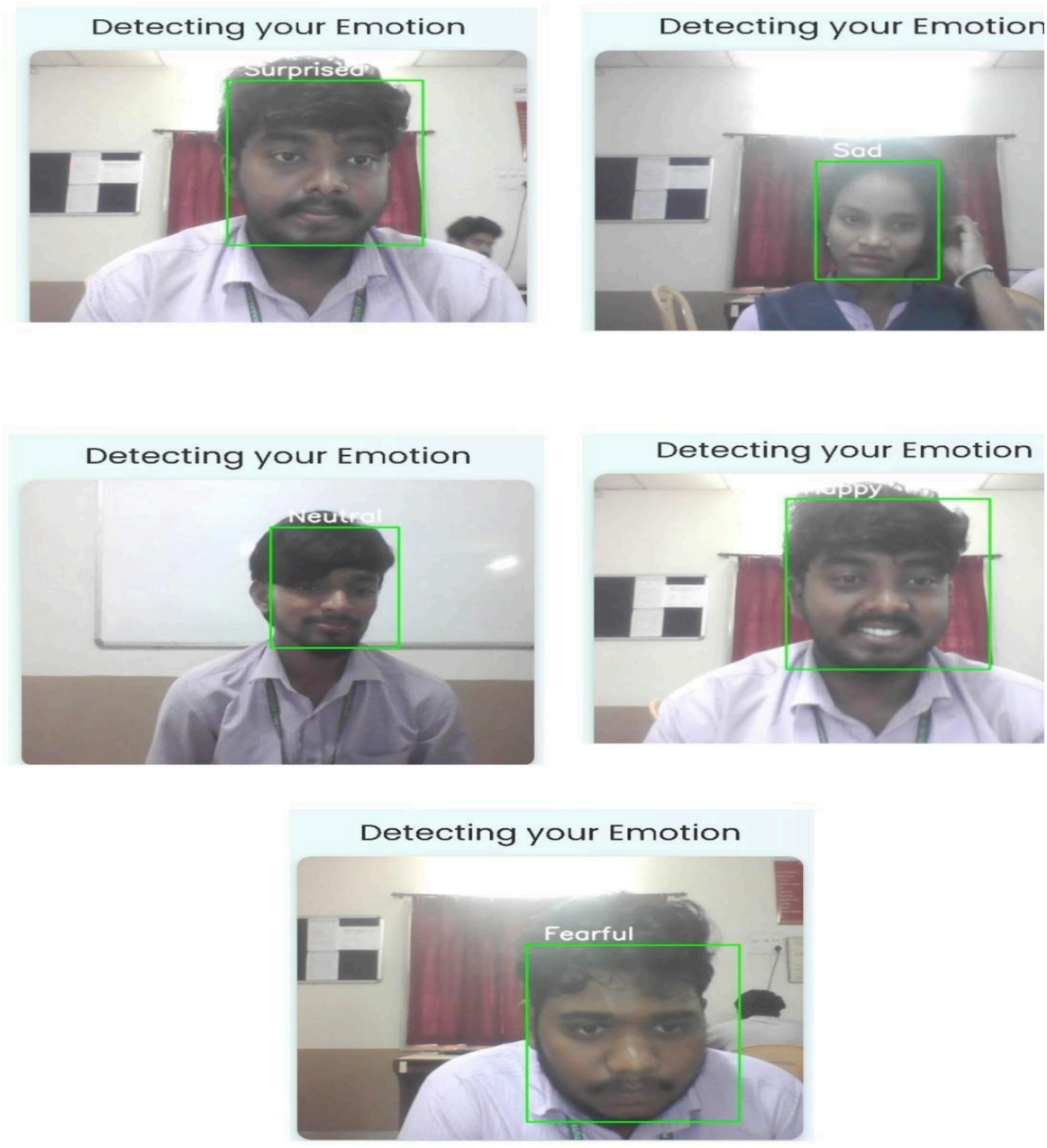


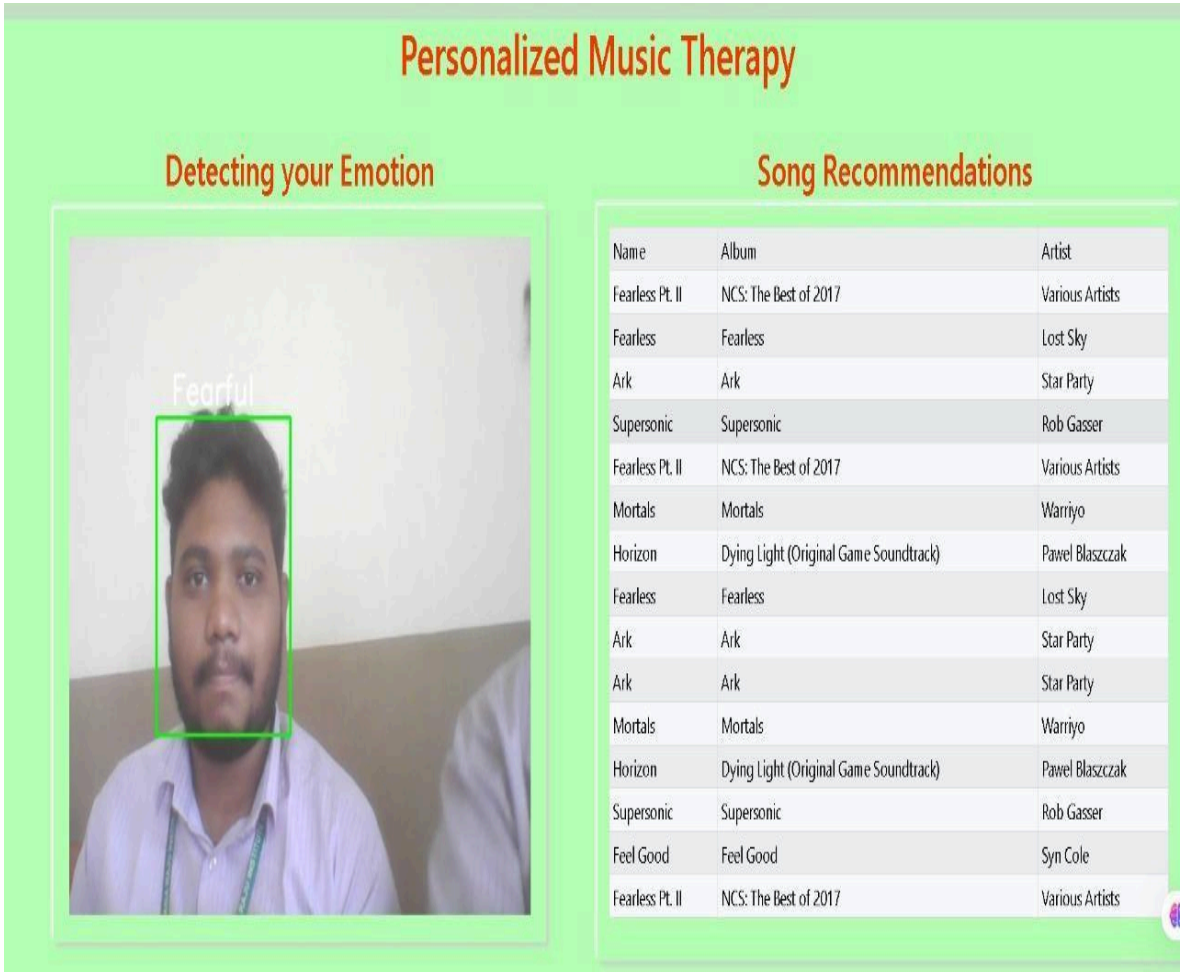
Fig: 2

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

Step 3: Music Recommendation with Spotify API

The detected emotion was passed to a dictionary (music_dist) mapping each emotion to a specific Spotify playlist link.

Using Spotipy, the system fetched tracks from the corresponding playlist via Spotify's Web API. The fetched tracks (Name, Album, Artist) were stored in a Pandas DataFrame and displayed or exported as CSV files for each emotion.



Name	Album	Artist
Fearless Pt. II	NCS: The Best of 2017	Various Artists
Fearless	Fearless	Lost Sky
Ark	Ark	Star Party
Supersonic	Supersonic	Rob Gasser
Fearless Pt. II	NCS: The Best of 2017	Various Artists
Mortals	Mortals	Warriyo
Horizon	Dying Light (Original Game Soundtrack)	Pawel Blaszczak
Fearless	Fearless	Lost Sky
Ark	Ark	Star Party
Ark	Ark	Star Party
Mortals	Mortals	Warriyo
Horizon	Dying Light (Original Game Soundtrack)	Pawel Blaszczak
Supersonic	Supersonic	Rob Gasser
Feel Good	Feel Good	Syn Cole
Fearless Pt. II	NCS: The Best of 2017	Various Artists

Fig: 3

The Journal of Computational Science and Engineering (TJCSE)
ISSN 2583-9055 (Media Online)
Vol 3, No 8, August 2025
PP 50–59

Step 4: Real-Time Playlist Switching
As the user's facial expression changed, the model continuously updated the detected emotion. If a new emotion was detected, the system automatically fetched and displayed a new set of music recommendations, allowing dynamic mood-based playlist switching.

To avoid excessive API calls or playlist flickering, emotion predictions were stabilized using a short buffer before switching.

Step 5: User Interface and Live Feedback
A Flask-based web interface was created using HTML and CSS with Neumorphism design, allowing users to interact visually with the application.
The interface displayed:

Current detected emotion
Suggested playlist from
Spotify
Track details like song name, album, and artist

Users could optionally click on songs to be redirected to Spotify for full playback. The application continued running until manually stopped by the user.

Conclusion

The Personalised Music Therapy system Based on Facial Expression is an intelligent application designed to enhance the music listening experience by dynamically selecting songs based on the user's emotional state. By leveraging facial expression recognition, machine learning, and music recommendation algorithms, the system personalizes playlists in real time, making music selection more intuitive and engaging. The project successfully integrates image processing, AI-based classification, and external APIs (such as Spotify) to provide an automated and user-friendly experience. The feedback mechanism further refines recommendations over time, ensuring a more accurate emotional-music mapping.

References

1. M. Keerthana, M. Shruthi, and S. Aravind Kumar, "Emotion Based Music Recommendation System," International Journal of Creative Research Thoughts (IJCRT), vol. 9, no. 6, pp. e356–e360, June 2021.

2. Anand R., Sabeenian R.S., Deepika Gurang, and Kirthika R., "AI-based Music Recommendation System using Deep Learning Algorithms," International Journal of Creative Research Thoughts (IJCRT), vol. 8, no. 5, pp. 1234–1240, May 2020.
3. Sriraj Katkuri, Mahitha Chegoor, K.C. Sreedhar, and M. Sathyanarayana, "Emotion Based Music Recommendation System," International Journal of Advanced Research in Computer and Communication Engineering, vol. 9, no. 5, pp. 45–50, May 2020.
4. Saurav Joshi, Tanuj Jain, and Nidhi Nair, "Emotion Based Music Recommendation System Using LSTM-CNN Architecture," International Journal of Creative Research Thoughts (IJCRT), vol. 11, no. 2, pp. 678–685, February 2023.
5. Tina Babu, Rekha R. Nair, and Geetha A., "Emotion-Aware Music Recommendation System: Enhancing User Experience Through Real-Time Emotional Context," arXiv preprint arXiv:2311.10796, November 2023.
6. Xinyu Chang, Xiangyu Zhang, Haoruo Zhang, and Yulu Ran, "Music Emotion Prediction Using Recurrent Neural Networks," arXiv preprint arXiv:2405.06747, May 2024.
7. Ramiz Mammadli, Huma Bilgin, and Ali Can Karaca, "Music Recommendation System based on Emotion, Age and Ethnicity," arXiv preprint arXiv:2212.04782, December 2022.
8. Erkang Jing, Yezheng Liu, Yidong Chai, Shuo Yu, Longshun Liu, Yuanchun Jiang, and Yang Wang, "Emotion-aware Personalized Music Recommendation with a Heterogeneity-aware Deep Bayesian Network," arXiv preprint arXiv:2406.14090, June 2024.
9. Florence and Uma, "Emotion-Based Music Recommendation System Using Facial Expression Analysis," International Journal of Research Publication and Reviews, vol. 5, no. 11, pp. 7876–7882, November 2024.
10. Patel et al., "Emotion-Based Music Player Using CNN for Facial Emotion Recognition," International Journal of Research Publication and Reviews, vol. 5, no. 11, pp. 7883–7890, November 2024.
11. M. Keerthana, M. Shruthi, and S. Aravind Kumar, "Emotion Based Music Recommendation System," International Journal of Creative Research Thoughts (IJCRT), vol. 9, no. 6, pp. e356–e360, June 2021.
12. Saurav Joshi, Tanuj Jain, and Nidhi Nair, "Emotion Based Music Recommendation System Using LSTM-CNN Architecture," International Journal of Creative Research Thoughts (IJCRT), vol. 11, no. 2, pp. 678–685, February 2023.