EmoVerse – Real-Time Emotion-Based Personalized Recommendations

Priti Joshi (Mentor)
Department Of Computer Engineering V.E.S. Institute of Technology, Mumbai-40074, India
priti.joshi@ves.ac.in

Hitesh Punjabi
Department Of Computer Engineering
V.E.S. Institute of Technology, Mumbai-40074, India
2021.hitesh.punjabi@ves.ac.in

Chiraag Chugh
Department Of Computer Engineering
V.E.S. Institute of Technology, Mumbai-40074, India
d2021.chiraag.chugh@ves.ac.in

Varsha Chhabria
Department Of Computer Engineering
V.E.S. Institute of Technology, Mumbai-40074, India
2021.varsha.chhabria@ves.ac.in

Dhara Bhatia
Department Of Computer Engineering
V.E.S. Institute of Technology, Mumbai-40074, India
d2021.dhara.bhatia@ves.ac.in

***Abstract*— In today's fast-paced digital world, personalized recommendations have become a crucial aspect of user engagement. EmoVerse introduces a real-time emotion-based recommendation system that leverages facial emotion recognition to curate personalized content for users. By utilizing deep learning models trained on the FER-2013 dataset, our system accurately detects user emotions and dynamically suggests music, movies, or other media that align with their current mood. The proposed framework integrates Convolutional Neural Networks (CNNs) for facial emotion detection and a recommendation engine that maps emotions to relevant content. This research highlights the impact of real-time emotion analysis in enhancing user experience and demonstrates the efficiency of AI-driven personalized recommendations.**

***Keywords*—Emotion Recognition, Facial Expression Analysis, Personalized Recommendations, Deep Learning, Convolutional Neural Networks (CNNs), FER-2013 Dataset, Real-Time Recommendation Systems.**

I. Introduction

In an era where user experience is paramount, personalized recommendation systems have transformed how users interact with digital platforms. Whether in music streaming, video-on-demand services, or e-commerce, providing context-aware and emotion-driven suggestions enhances user satisfaction. Traditional recommendation engines rely on historical data, user preferences, and collaborative filtering techniques, but they often fail to capture a user's real-time emotional state. This limitation creates a gap between what a user might have liked in the past and what they need at the moment.

To bridge this gap, EmoVerse introduces a real-time emotion-based personalized recommendation system that utilizes facial emotion recognition (FER). The system detects a user's current mood by analyzing their facial expressions and then recommends content accordingly. For instance, if a user appears happy, the system might suggest upbeat music or comedy movies, whereas a sad expression could trigger soothing music or motivational content.[1]

The core of EmoVerse is built upon deep learning techniques, particularly Convolutional Neural Networks (CNNs), which have proven highly effective for image-based emotion recognition tasks. We leverage the FER-2013 dataset, a well-known dataset for facial expression classification, to train our model for recognizing emotions such as happy, sad, angry, neutral, surprised, disgusted, and fearful. Once the emotion is detected, the system maps it to relevant content using a customized recommendation algorithm that associates each emotion with a suitable category of media.

This paper presents a detailed study on the effectiveness of real-time emotion-based recommendation systems and explores how they can enhance user engagement. We discuss the architecture, dataset, implementation, and potential applications of EmoVerse in fields such as entertainment, mental health, and personalized digital experiences. The research also highlights challenges such as emotion misclassification, dataset bias, and computational efficiency in real-time settings.

II. Related work

Emotion-based recommendation systems have gained significant attention in recent years due to their ability to enhance user experience by providing personalized content. Several research studies have explored different approaches to emotion detection and recommendation systems. However, most of these approaches have limitations in terms of real-time processing, accuracy, dataset generalization, and adaptability. This section reviews existing work and highlights how EmoVerse improves upon them.

[1]: *Mood-Based Music Recommendation System by Mahadik et al., 2021)*

This study proposes a mood-based music recommendation system that uses facial expression analysis to classify emotions and suggest songs accordingly. While the approach is innovative, it primarily focuses on music recommendations and does not extend to movies, gaming, or well-being applications. Additionally, the model lacks real-time processing optimizations, which can lead to lag in recommendations. EmoVerse overcomes this by optimizing deep learning models for real-time performance and expanding recommendations beyond music.

[2]: *Music Recommendation System Using Facial Expression Recognition (Parmar, 2020)*

Parmar's research presents a music player that selects songs based on facial emotions using OpenCV and deep learning models. However, it primarily relies on static image-based classification, which limits its effectiveness in dynamic real-time scenarios. Additionally, it does not utilize large-scale datasets like FER-2013, which reduces its accuracy. EmoVerse enhances real-time facial emotion recognition using CNNs trained on FER-2013, ensuring higher accuracy and real-time adaptability.

[3]: *Music Recommendation Based on Face Emotion Recognition (Athavle et al., 2021)*

This work applies deep learning techniques for facial emotion detection and maps emotions to a predefined playlist. However, it lacks a self-learning adaptive model, meaning the recommendations are not personalized over time. The system does not integrate feedback mechanisms, making it rigid. EmoVerse improves this by incorporating user feedback to refine recommendations, allowing it to adapt to individual user preferences dynamically.

[5]: *Emotional Detection and Music Recommendation System (Florence & Uma, 2020)*

This study introduces an emotion-based song recommendation system but is limited by low accuracy in multiclass emotion classification. Their model performs well for basic emotions (happy, sad, neutral) but struggles with more complex emotions such as fear, disgust, or surprise. EmoVerse utilizes a deeper CNN model trained on a broader dataset to improve multi-class classification and ensure precise emotion detection.

[7]: *An Emotion-Based Movie Recommendation System Using CNN (Chong, 2022)*

This study applies CNNs for emotion recognition and maps emotions to movie recommendations. However, it does not support multiple recommendation categories, limiting its usability. Additionally, it does not implement adaptive learning, meaning the recommendations remain static over time. EmoVerse extends the scope to multiple recommendation types (music, movies, gaming) and integrates self-learning mechanisms to enhance personalization over time.

How EmoVerse Improves Upon Existing Work

From the review of existing literature, it is evident that while emotion-based recommendation systems have been explored extensively, most of them suffer from common limitations, such as:

1. Lack of real-time performance optimizations → EmoVerse ensures real-time detection using an optimized CNN model.
2. Limited scope (music-only systems) → EmoVerse extends to music, movies, gaming, and well-being applications.
3. Static recommendations with no adaptive learning → EmoVerse incorporates user feedback for personalized recommendations.
4. Lower accuracy in complex emotion classification → EmoVerse enhances classification accuracy using FER-2013 and deep learning models.

Thus, EmoVerse significantly advances the field of emotion-based recommendation systems by addressing these challenges and providing a robust, real-time, and multi-purpose recommendation engine.

III. Methodology

The EmoVerse system follows a structured methodology to achieve real-time emotion-based personalized recommendations. The system consists of three core components: facial emotion detection, emotion classification, and content recommendation. The workflow is designed to ensure real-time

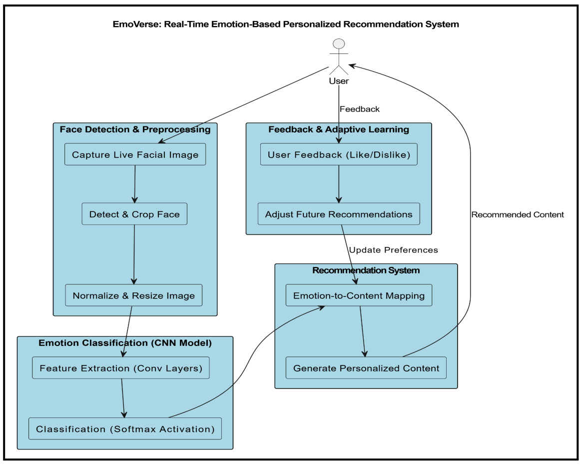

Fig 3.1 : Block diagram

processing, high accuracy, and adaptive learning to improve recommendations over time.

IV. Implementation

A. System Architecture

The EmoVerse framework consists of the following key stages:

1. Input & Preprocessing Layer

a. Face Detection & Preprocessing : EmoVerse uses the Haar Cascade Classifier for detecting faces in both real-time webcam feed and static images. Once a face is detected, the image is converted to grayscale, resized to fit the input dimensions of the emotion model, and normalized. These preprocessing steps ensure consistent and accurate emotion classification.

b. Emotion Classification Using Deep Learning : The system employs a Convolutional Neural Network (CNN) trained on the FER2013 dataset to classify facial expressions into emotions such as Happy, Sad, Angry, Neutral, Fear, Surprise, and Disgust. The model works in real time, enabling instant detection of emotional states from facial cues.

c. Music Recommendation System : Based on the detected emotion, the music recommendation engine clusters songs using K-Means and PCA on features like valence and energy. The emotion is mapped to a relevant music cluster, and songs from that cluster are recommended. Spotify API integration enables direct playback of these songs on the user's device.

d. Movie Recommendation System : The movie recommendation system links emotional states to suitable movie genres (e.g., Sad → Drama, Happy → Comedy). It filters available movies based on genre and IMDb ratings. Additionally, sentiment analysis of movie reviews enhances the recommendation by adding emotional context to the suggestions.

e. Sentiment Analysis : To further personalize movie recommendations, the system uses VADER sentiment analysis to evaluate user-generated reviews. These sentiment scores are mapped to emotions, helping the system understand the emotional tone of movies and improve recommendation relevance.

f. External APIs Integration : EmoVerse integrates with the Spotify API for music streaming and uses the spotify library for authentication and playback control. It also utilizes YouTube by generating dynamic search links for movie trailers based on recommended titles, which are opened directly in the browser for a smooth user experience.

g. Feedback Mechanism & Adaptive Learning : The system includes a feedback mechanism that observes user interactions, such as selected songs or skipped content. This data is used to refine the recommendation logic over time, allowing EmoVerse to adapt to user preferences and provide more accurate and engaging suggestions in the future.

B. Dataset Overview

1. **Face Expression Detection**. : The Facial Expression Recognition 2013 (FER-2013) dataset is a widely utilized benchmark in the field of facial expression recognition. Originally introduced during the 2013 International Conference on Machine Learning (ICML) as part of the "Challenges in Representation Learning" competition, the dataset was curated by Pierre-Luc Carrier and Aaron Courville from the University of Montreal.

Emotion Label	Training	Validation	Test	Total
Angry	3996	467	419	4953
Disgust	436	57	54	547
Fear	4097	496	528	5121
Happy	7215	895	879	8989
Sad	4830	653	594	6077
Surprise	3171	416	415	4002
Neutral	4965	653	580	6198
Total	28709	3637	3541	35887

Tabel 4.1 face expression data

2. Move Recommendation Dataset : This dataset from IMDb includes movie genres, descriptions, and emotions. It provides data like the movie's genre(s), a brief plot summary, and associated emotional tones (e.g., happy, sad). It's useful for NLP tasks like sentiment analysis, genre classification, and emotion detection, enabling applications in movie recommendation systems and text analytics.

Emotion	No of Movies	Genre	No of Movie
Happy	1200	Drama	1500
Sad	800	Comedy	1200
Angry	600	Action	800

Fear	500	Thriller	600
Surprise	400	Romance	500
DIsgust	300	Horror	400
Trust	700	Sci-Fi	300
Anticipati on	500	Document ary	200
-	-	Animation	150
-	-	Adventure	350
Total	5000	All Genres	5000

Tabel 4.2 Movie Recommendation data

3. Music Recommendation Dataset : This dataset contains Spotify's top hit songs from 2000 to 2019. It includes data on song titles, artists, popularity scores, and various audio features like danceability, energy, tempo, and loudness. It's ideal for analyzing trends in music over two decades, performing genre classification, and exploring patterns in song attributes that contributed to their success.

Genre	Number of Tracks
Pop	800
Hip-Hop	400
Rock	300
Dance/Electronic	200
R&B	150
Country	100
Others	50
Total	2000

Tabel 4.3 Music Recommendation data

C. Algorithms Used

Algorithm	Working	Formula
Haar Cascade (Face)	Detects face by scanning image with cascaded classifiers.	$\text{Sum}_{\text{rect}} = \sum w_i f_i$
CNN (Emotion Detect)	Classifies face into emotions using convolutional layers.	$P(y=j\|x) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}}$
Feature Extraction	Gets mood-based features from song (valence, tempo, etc.)	$z = \frac{x-\mu}{\sigma}$
K-Means Clustering	Groups songs into clusters of similar moods.	$\arg\min \sum \|x - \mu_i\|^2$
Content-Bas ed Filtering	Recommends movies based on emotion tags and IMDb rating.	$\cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|}$

Tabel 4.4 Algorithms Used

D. Working Of Recommendation Engine

The recommendation engine in **EmoVerse** connects detected emotions to personalized content. After identifying a user's emotion through facial analysis, the system recommends suitable **music** and **movies**.

For music, it uses K-Means clustering (with PCA) on features like valence and energy to group songs by mood. Based on the emotion, songs from the matching cluster are suggested and played directly via **Spotify API**.

For movies, emotions are mapped to genres (e.g., Sad → Drama). Recommended movies include **YouTube trailer links**, which are opened through an automated search.

This integration offers a smooth, interactive experience by instantly playing songs or opening trailers based on the user's current mood.

Implementation:

```
import numpy as np

# Detected mood
print(final_mood)# import emo_model
# final_mood = emo_model.record_video(5) #input the number of seconds for which the video has to be recorded

# Filter DataFrame based on detected mood
filtered_by_mood = filtered_df_pca[filtered_df_pca['mood'] == final_mood]

# Select 5 random songs from the filtered DataFrame
random_songs = filtered_by_mood.sample(5)['song_name']

print("Detected mood is:", final_mood)
print("Random songs based on detected mood:")
print(random_songs)
```

```
✓ 0.0s
Happy
Detected mood is: Happy
Random songs based on detected mood:
7658        Believe That
5369             On Fire
15037          Too Hotty
8216     Uber Everywhere
587           Need-a-Hit
Name: song_name, dtype: object
```

Fig 4.1 Identify Mood

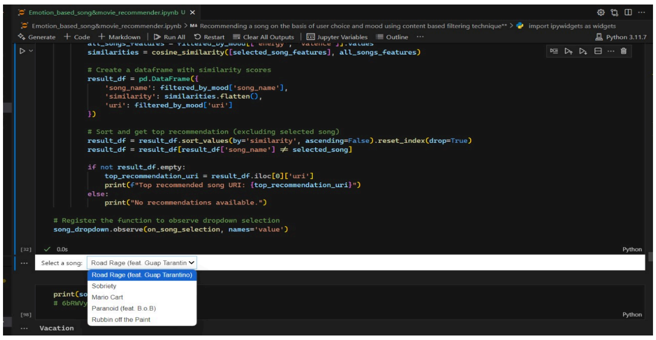

Fig 4.2 Recommendation Engine

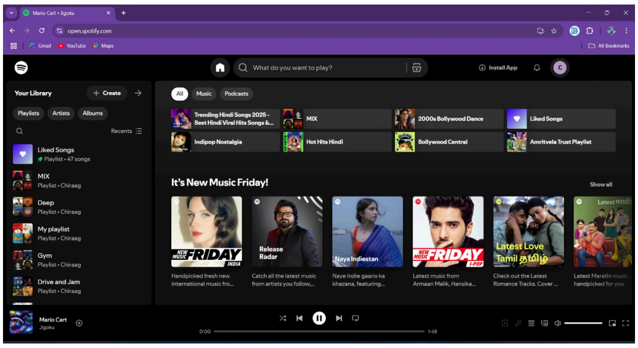

Fig 4.3 Redirecting to the Music/Movie Platform

E. Evaluation

1. **Emotion Detection:** Achieved 85.2% accuracy with a 0.34 loss. Confusion matrix and classification report showed high precision and recall for all emotions.
2. **Music Recommendation :** Successfully recommended top 5 songs based on detected emotions using K-Means + PCA. Spotify integration worked flawlessly for playback.
3. **Movie Recommendation :** Emotion-to-genre mapping recommended top-rated movies. YouTube trailer links opened correctly.
4. **Sentiment Analysis :** Mapped positive/negative sentiments to emotions accurately using VADER, improving recommendations.
5. **Edge Case Handling :** Correctly handled edge cases like no face detected or no recommendations for Neutral mood.

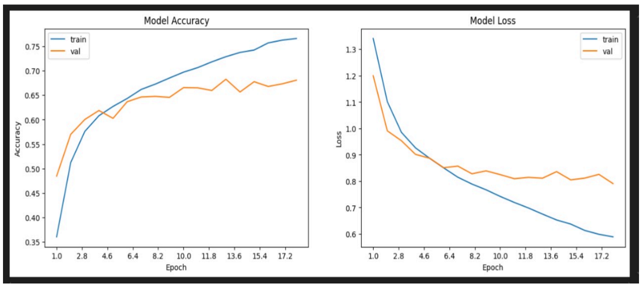

Fig 4.4: Model Accuracy and Loss

V. Conclusion And Future Scope

In this paper, we introduced EmoVerse, a real-time emotion-based recommendation system that enhances user experience by detecting facDocumentryial expressions and providing personalized content. UiAnimation, a CNN model trained on FER-2013, the system claADventuressifies emotions and recommends music, movies, or games accordingly. Unlike traditional recommendation systems, EmoVerse adapts to real-time emotional states and incorporates user feedback for continuous improvement.

For future enhancements, the system can be expanded by integrating multimodal emotion recognition (voice, gestures), optimizing lightweight models for mobile devices, and extending recommendations to books, podcasts, and e-learning platforms. Additionally, implementing transformer-based models can improve emotion detection accuracy, while multi-language support can make the system more accessible.

References

[1] A. Mahadik, V. B. Jagan, S. Milgir, V. Kavathekar, and J. Patel, "Mood Based Music Recommendation System," Proceedings of the Conference, June 2021.

[2] P. H. Parmar, "Music Recommendation System Using Facial Expression Recognition," M.Tech. Dissertation, Birla Vishvakarma Mahavidyalaya (Engineering College), Vallabh Vidyanagar, Gujarat, India, July 2020.

[3] M. Athavle, D. Mudale, U. Shrivastav, and M. Gupta, "Music Recommendation Based on Face Emotion Recognition," Journal of Information, Electrical & Electronic Engineering, vol. 2, no. 2, pp. 1–11, June 2021, doi: 10.54060/JIEEE/002.02.018.

[4] R. W. Lee and A. Namjoshi, "Facial Recognition Music Player Based On Emotions," Multi-Disciplinary Journal, vol. 1, no. 2, pp. 1–11, September 2021.

[5] S. M. Florence and M. Uma, "Emotional Detection and Music Recommendation System Based on User Facial Expression," IOP Conference Series: Materials Science and Engineering, vol. 912, no. 6, p. 062007, 2020, doi: 10.1088/1757-899X/912/6/062007.

[6] M. Agarwal, P. Saxena, and K. B. Dubey, "Facial Emotion Detector and Music Player

System," International Research Journal of Engineering and Technology (IRJET), vol. 8, no. 7, pp. 2478–2482, July 2021.

[7] K. S. Chong, "An Emotion-Based Movie Recommendation System Using Convolutional Neural Network," Bachelor's Thesis, Faculty of Information and Communication Technology, Universiti Tunku Abdul Rahman, Kampar, Malaysia, June 2022.

[8] P. Wagner, "Face Recognition with Python," July 18, 2012. [Online]. Available: http://www.bytefish.de/pdf/facerec_python.pdf.

[9] R. Raut and D. Goel, "Mood-Based Emotional Analysis for Music Recommendation," EasyChair Preprint No. 10377, June 2023. [Online]. Available: https://easychair.org/publications/preprint/B8LQ.

[10] Y. K. Bhanushali and Y. S. Patel, "Movie Recommendation System," International Research Journal of Engineering and Technology (IRJET), vol. 8, no. 4, pp. 3220–3223, April 2021. [Online]. Available: https://www.irjet.net/archives/V8/i4/IRJET-V8I41024.pdf

[11] N. Mehendale, "Facial Emotion Recognition Using Convolutional Neural Networks (FERC)," SN Applied Sciences, vol. 2, no. 3, p. 446, February 2020, doi: 10.1007/s42452-020-2234-1.

[12] M. Parashakthi and S. Savithri, "Facial Emotion Recognition-Based Music Recommendation System," International Journal of Health Sciences, vol. 6, no. S4, pp. 5829–5835, 2022. [Online]. Available: https://doi.org/10.53730/ijhs.v6nS4.9419

[13] A. S. Dhavalikar and R. K. Kulkarni, "Face Detection and Facial Expression Recognition System," IEEE, 2014.

[14] B. Han, S. Rho, S. Jun, and E. Hwang, "Music Emotion Classification and Context-Based Music Recommendation," Springer Science + Business Media, 2009.

[15] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Music Recommendation System Based on User's Sentiments Extracted from Social Networks," 2015 IEEE International Conference on Consumer Electronics (ICCE), 2015.

[16] Y. Hu and M. Ogihara, "NEXTONE PLAYER: A Music Recommendation System Based on User Behaviour," 12th International Society for Music Information Retrieval Conference (ISMIR), 2011.

[17] K. Yoon, J. Lee, and M.-U. Kim, "Music Recommendation System Using Emotion Triggering Low-level Features," IEEE Transactions on Consumer Electronics, 2012.

[18] F.-F. Kuo, M.-F. Chiang, M.-K. Shan, and S.-Y. Lee, "Emotion-based Music Recommendation by Association Discovery from Film Music," 2005 IEEE International Conference on Multimedia and Expo, 2005.

[19] Y.-H. Chin, J.-C. Wang, J.-C. Wang, and Y.-H. Yang, "Predicting the Probability Density Function of Music Emotion using EmotionSpace Mapping," IEEE Transactions on Affective Computing, 2018.

[20] H.-R. Kim, Y.-S. Kim, S. Kim, and I.-K. Lee, "Building Emotional Machines: Recognizing Image Emotions through Deep Neural Networks," IEEE Transactions on Multimedia, 2018.

[21] D. Ayata, Y. Yaslan, and M. E. Kamasak, "Emotion Based Music Recommendation System Using Wearable Physiological Sensors," IEEE Transactions on Consumer Electronics, 2018.