iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56769

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

FeelFlow: Music Recommendation System Based on your Facial Expression

Prof. Vandana Navale[1], Sanjana V. Phand[2], Pranav K. Patil[3], Dhananjay V. Rangat[4], Rohan J. Sathe[5]
[1, 2, 3, 4, 5]*Department of Computer Engineering All India Shri Shivaji Memorial Society's College of Engineering, Pune*

Abstract: ***This study introduces an innovative Emotion-Aware Music Recommender System that leverages facial feature analysis for real-time emotion detection, aiming to elevate the personalization and user engagement in music recommendations. Through the integration of computer vision and machine learning techniques, the system interprets facial expressions to discern user's emotional states before recommending music to them. Integrating advanced deep learning and evolutionary techniques improves the accuracy and efficiency of facial expression detection. Implementation of the proposed recommender system is performed using Viola-Jones algorithm, CNN, Principal Component Analysis (PCA) techniques. Experimental results confirm the effectiveness of this method and demonstrate its potential to improve music recommendations.***
Keywords: Music Recommender System (MRS), Emotion Detection, Deep Learning, Convolutional neural networks (CNN), Viola-Jones Algorithm, Principal Component Analysis (PCA), Facial Feature Analysis, Deep Neural Network (DNN)

I. INTRODUCTION

In recent years, the field of music recommendation systems has witnessed a paradigm shift towards enhancing user experience through increased personalization. Traditional recommendation engines primarily rely on user preferences, historical data, and collaborative filtering to generate music suggestions. However, these approaches often overlook a crucial aspect of user engagement – emotions. Music, as a profoundly emotive and subjective art form, elicits a wide range of emotional responses from listeners. This review paper explores the integration of emotion detection through facial feature analysis in MRS, aiming to provide a comprehensive understanding of the emerging landscape in this domain. Recognizing the significance of emotions in shaping individual musical preferences, researchers have increasingly turned to advanced technologies such as computer vision and machine learning to capture and interpret users' emotional states in real-time. In the rapidly evolving landscape of MRS, the integration of facial expression analysis has emerged as a transformative avenue, enriching user experiences through personalized recommendations.

Leveraging advanced algorithms such as Viola-Jones, CNN and SVM, researchers have sought to capture the nuanced emotional responses of users during music consumption, thereby enhancing the precision and adaptability of recommendation engines.

The Viola-Jones algorithm, renowned for its efficiency in real-time object detection, has found application in the preliminary stages of facial feature extraction.

By efficiently identifying key facial landmarks and expressions, Viola-Jones serves as a foundational step in the extraction process, laying the groundwork for subsequent analysis. Following this initial extraction, CNN come into play, demonstrating their prowess in image recognition tasks. CNNs excel at learning hierarchical representations, making them well-suited for discerning intricate patterns within facial expressions. In the context of Music Recommender Systems, CNNs contribute to the nuanced understanding of users' emotional states, allowing for a more granular and accurate emotional profile construction. Complementing the feature extraction and representation learning capabilities of Viola-Jones and CNNs, SVM step in as powerful classifiers.

SVMs excel in discerning complex relationships within datasets, making them invaluable for mapping facial expressions to corresponding emotional states. By training SVM models on labelled datasets of facial expressions and associated emotions, the system refines its ability to predict user emotions, thereby facilitating a more precise alignment of music recommendations with the users' emotional context.

This paper navigates through the multifaceted landscape of these algorithms, exploring their individual roles and synergies in the development of an Emotion-Driven Music Recommender System. As we delve into the applications of Viola-Jones, CNN, and SVM in the context of facial expression analysis for music recommendations, we aim to elucidate the collective impact of these algorithms in shaping the future of personalized and emotionally resonant music discovery experiences.

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

The study elucidates the utilization of facial features to detect emotions, enabling a personalized music recommendation system tailored to individual emotional states. The Machine Learning Algorithms used are Viola-Jones for HAAR features and feature extraction, PCA, CNN, SVM for feature vectors and multi class Adaboost during dynamic time warping.[1]. The study highlights the implementation of machine learning techniques to develop a sophisticated system capable of retrieving and recommending music, catering to individual preferences and enhancing the user's music discovery experience. we present a music retrieval and recommendation system using machine learning techniques. We propose a DNN based note transcription method and create a complete query by humming music retrieval system, which we test using the standard MIREX Query by humming data set.[5].

The study presents findings emphasizing the development of a recommendation system that utilizes emotions to personalize music suggestions, enhancing user experiences based on emotional responses. Using SVM, CNN, Mel spectrogram.[6]. The study reveals insights into leveraging facial expressions for music therapy, highlighting the potential of using emotions detected from facial cues to personalize therapeutic musical experiences. A user-friendly intelligent system employing machine learning empowers therapists by effortlessly detecting emotions through a camera and tailoring music accordingly, emphasizing simplicity and intuitive music listening experiences.[9]. The study shows use of Artificial Intelligence, Deep Learning, Deep Reinforcement Learning Emotional health, Neural Activity, Neural Networks, Recurrent Neural Networks.[7]

Table 1
LITERATURE SURVEY

Year	Reference no.	Key Points	Advantages	Disadvantages
2019	1	Usage of Viola-Jones, PCA, CNN, SVM.	Elucidates the utilization of facial features to detect emotions.	Limited emotion detection accuracy. Small emotion category set. Potential mismatch with individual music preferences.
2021	5	Proposed a Deep Neural Network (DNN) based note transcription method	Recommending music, catering to individual preferences. Encouraging Results: QBH system's promising performance.	While the system demonstrates success, it's important to compare its results with commercial techniques used by platforms like Spotify and Amazon Music to understand its competitiveness further.
2021	6	Use of Mel Spectogram, Support Vector Machines and CNN.	Effective Content-Based Strategy: The system leverages the Big Five psychological model, user profiles, and mood to create a highly effective browsing experience for audio collections.	Challenges with Poor Musical Chronology: The system may face challenges when dealing with users with limited or incomplete music listening history.
2020	7	Leveraging facial expressions for music therapy	The system is able to generate personalized music therapy by adapting.	The system in some cases needs to be accompanied by other kinds of treatments.
2020	9	Application of AI, Deep Learning, Neural Activity, Neural Networks	Fisher face recognition method proves to be efficient because it works better with additional features.	Limited Model Details: Lack of information on the model and its training process. Privacy Concerns

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

III. MACHINE LEARNING ALGORITHMS

A. Viola - Jones Algorithm in Music Recommender System:

The Viola-Jones algorithm has emerged as a pivotal tool in the realm of Music Recommender Systems (MRS), specifically in the context of leveraging facial expressions for enhancing user experience. Widely recognized for its efficiency in real-time object detection, the Viola-Jones algorithm is strategically employed in the preliminary stages of facial feature extraction within the MRS framework. In the context of music recommendation, the Viola-Jones algorithm plays a critical role in swiftly and accurately identifying key facial landmarks and expressions. By employing a cascading series of classifiers, Viola-Jones efficiently sifts through image data, pinpointing regions of interest that correspond to distinct facial features indicative of various emotional states. The utility of Viola-Jones lies in its ability to robustly capture facial expressions in real-time, facilitating the continuous monitoring of users' emotional responses to the music. This dynamic extraction of facial features enables the Music Recommender System to adapt and respond promptly to shifts in the user's emotional state, thereby enhancing the system's capacity to generate timely and emotionally resonant music recommendations.

B. Convolutional Neural Networks (CNNs) in Music Recommender System:

Convolutional Neural Networks (CNNs) are very important components in the augmentation of Music Recommender Systems (MRS), particularly in the realm of leveraging facial expressions to enhance user-centric recommendations. The versatility of CNNs in image recognition and feature extraction makes them an invaluable asset in discerning intricate patterns within facial expressions, thereby contributing to the depth and precision of emotional profiling within MRS. In the context of music recommendation, CNNs play a crucial role in capturing the nuanced details of facial expressions, allowing for a more granular understanding of users' emotional states. By employing convolutional layers that automatically learn hierarchical representations, CNNs excel at extracting complex features from facial imagery. This capability is paramount in decoding the subtleties of emotional cues conveyed through various expressions, enabling the Music Recommender System to create a more sophisticated emotional profile for each user. The hierarchical feature extraction facilitated by CNNs enables the system to identify and interpret subtle variations in facial expressions, enriching the emotional context of the user experience. By leveraging the deep learning capabilities of CNNs, MRS can unlock new dimensions of user engagement, offering more tailored and emotionally relevant music suggestions. As we navigate the intricate interplay between CNNs and Music Recommender Systems, this exploration aims to contribute to the evolving landscape of personalized, emotion-driven music discovery experiences.

C. Support Vector Machines (SVM) in Music Recommender System:

Support Vector Machines (SVMs) have become instrumental in the evolution of Music Recommender Systems (MRS), particularly in the context of leveraging facial expressions to enhance the personalization and emotional resonance of music recommendations. SVMs, known for their proficiency in complex pattern recognition and classification tasks, play a critical role in mapping facial expressions to corresponding emotional states within the MRS framework. In the realm of music recommendation, SVMs are employed as robust classifiers to discern the intricate relationships between facial features and the associated emotional responses. Trained on labelled datasets that correlate facial expressions with specific emotions, SVMs enable the system to learn and generalize these relationships, fostering a more nuanced understanding of users' emotional states during music consumption. The application of SVMs in Music Recommender Systems contributes significantly to the refinement of emotional profiling. By effectively classifying facial expressions into distinct emotional categories, SVMs enhance the system's ability to tailor music recommendations to the user's current emotional context. This classification process allows for a more precise alignment between the recommended music and the user's emotional state, thereby elevating the overall quality of the music discovery experience. SVMs contribute to the adaptive nature of the system, allowing it to continuously learn and improve its ability to decipher and respond to users' evolving emotional states. By navigating the synergies between SVMs and Music Recommender Systems, this exploration contributes to the ongoing enhancement of personalized, emotion-centric music discovery experiences.

D. Principal Component Analysis (PCA) in Music Recommender System:

Principal Component Analysis (PCA) has emerged as a valuable technique in Music Recommender Systems (MRS), specifically when integrating facial expressions for enhanced personalization. PCA, a dimensionality reduction method, plays a crucial role in extracting meaningful features from facial expression data within the MRS framework. In the context of music recommendation, PCA serves as a tool to reduce the complexity of facial expression datasets while retaining essential information.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

By transforming the high-dimensional space of facial features into a lower-dimensional subspace, PCA enables the system to focus on the most salient components of the data. This reduction in dimensionality not only streamlines computational processes but also facilitates a more efficient and effective representation of the underlying emotional states expressed through facial cues. The application of PCA in Music Recommender Systems contributes to the creation of a concise and informative emotional profile for users. By capturing the most significant variations in facial expressions, PCA refines the system's ability to understand and respond to users' emotional responses during music consumption. This streamlined representation facilitates a more agile and adaptive recommendation engine, as it hones in on the essential features that drive users' emotional engagement with music. PCA, through its dimensionality reduction capabilities, enriches the emotional profiling process, fostering a more streamlined and nuanced understanding of users' emotional experiences. By navigating the synergies between PCA and Music Recommender Systems, this exploration contributes to the ongoing refinement of personalized, emotion-centric music discovery experiences.

IV. COMPONENTS OF RECOMMENDATION SYSTEM

A. *Data Collection*

Data collection involves gathering a diverse set of facial expression and music-related datasets for training and testing the Emotion Detection and Music Recommendation Models, respectively. The chosen dataset FER2013 for facial expressions and Million Song Dataset for music features, play a crucial role in model development. Ensuring tdatasets encompass a wide range of emotions and music preferences enhances the system's ability to generalize. Ethical considerations regarding user privacy and consent are integral during the data collection process. This comprehensive dataset forms the foundation for building accurate and inclusive models in the music recommendation system.

B. *Data Preprocessing*

Data preprocessing is a critical step in preparing the collected data for effective model training in the music recommendation system. This involves several techniques as following:

1) .Data Augmentation: Data augmentation involves diversifying a dataset without changing its inherent meaning. In the context of facial expression analysis, techniques like rotation, scaling, and flipping are applied to the images to increase variability and enhance the model's ability to generalize to different facial expressions. This process aids in improving the robustness and performance of the Emotion Detection Model.
2) Data Normalization: Data Normalization is a crucial step in the music recommendation system's Emotion Detection Model. It involves scaling the pixel values of facial expression images to a standardized range, often between 0 and 1. This process ensures consistent data input to the CNN during training and helps stabilize the learning process. Normalization contributes to improved model convergence and efficiency, enhancing the overall effectiveness of emotion detection based on facial expressions.

C. *Exploratory Data Analysis*

Data exploration in this system entails understanding and analysing the facial expression dataset, including FER2013, CK+, or AffectNet. It involves investigating the distribution of emotions, visualizing sample images, and gaining insights into the dataset's characteristics. This exploration phase guides decisions in data preprocessing and model design, ensuring a comprehensive understanding of the data before model development

D. *Model Design and Training*

Model building in this system involves designing and training two main components: the Emotion Detection Model and the Music Recommendation Model. For the Emotion Detection Model, facial expression datasets like FER2013 are prepared with augmentation. A CNN architecture, potentially leveraging transfer learning, is chosen, and training involves selecting appropriate loss functions and optimizers.

Evaluation metrics such as accuracy guide the fine-tuning process. The Music Recommendation Model incorporates audio feature datasets and user preferences, using either collaborative or content-based filtering. Training and evaluation follow similar steps, considering loss functions and optimizers. The final integration combines both models to create a cohesive system for real-time processing and recommendation. The development process prioritizes thorough testing, user feedback, and validation to ensure optimal performance and user satisfaction.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Model Evaluation

Model evaluation in this music recommendation system involves assessing the Emotion Detection Model and Music Recommendation Model's performance. For the Emotion Detection Model, metrics such as accuracy, precision, recall, and F1-score measure its ability to accurately identify facial expressions. The Music Recommendation Model is evaluated using metrics like mean squared error or ranking metrics to gauge the accuracy of personalized playlist recommendations. Fine-tuning follows, adjusting hyperparameters based on evaluation results to optimize model performance. Thorough evaluation ensures the reliability and effectiveness of the entire recommendation system.

F. Deployment

The user interface in this music recommendation system is designed to provide an intuitive and user-friendly interaction platform. It includes features for capturing real-time facial expressions, enabling users to easily navigate and access personalized playlists. The interface incorporates elements for clear feedback on detected emotions and recommended music choices, enhancing the overall user experience.

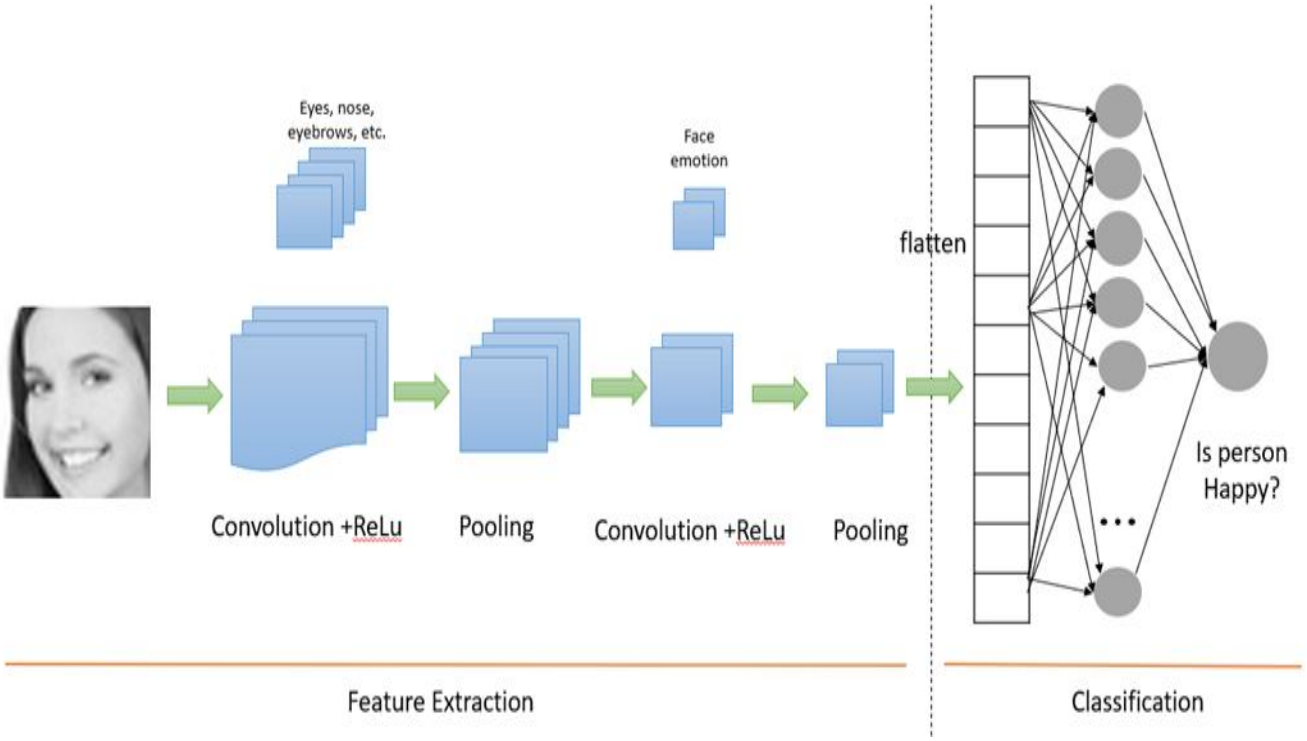


Fig.1. Convolutional Neural Network

V. METHODOLOGY

The methodology for developing a music recommendation system based on facial expressions employs a multi-faceted approach. Initial steps involve acquiring and preprocessing datasets, including facial expression datasets like FER2013 and music datasets such as Million Song Dataset. The Emotion Detection Model is built using a CNN architecture with Rectified Linear Unit (ReLU) activation functions for feature map extraction and softmax activation for emotion classification. The backend server is set up using TensorFlow Serving and FastAPI, ensuring efficient model deployment. To optimize the model, quantization techniques are applied, and TensorFlow Lite is explored for deploying lightweight models suitable for mobile platforms.

User-friendly interfaces are developed using React.js for the web and React Native for mobile, and the application is deployed on Google Cloud Platform (GCP) for scalability. Feature extraction and classification mechanisms are implemented in the CNN, incorporating pooling layers for down-sampling feature maps. Evaluation metrics such as accuracy, precision, and recall are utilized for model assessment, with continuous improvement facilitated through a feedback loop based on user interactions and preferences. Overall, this methodology integrates cutting-edge techniques and tools to create a robust and user-centric music recommendation system.

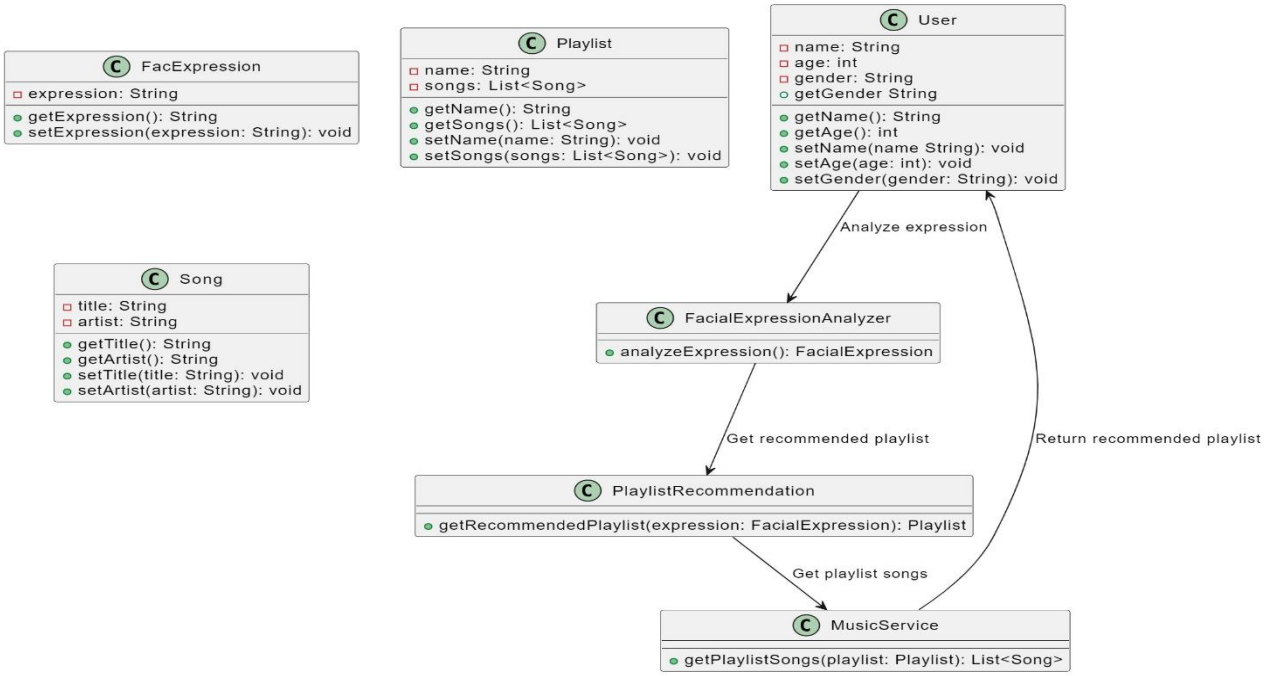


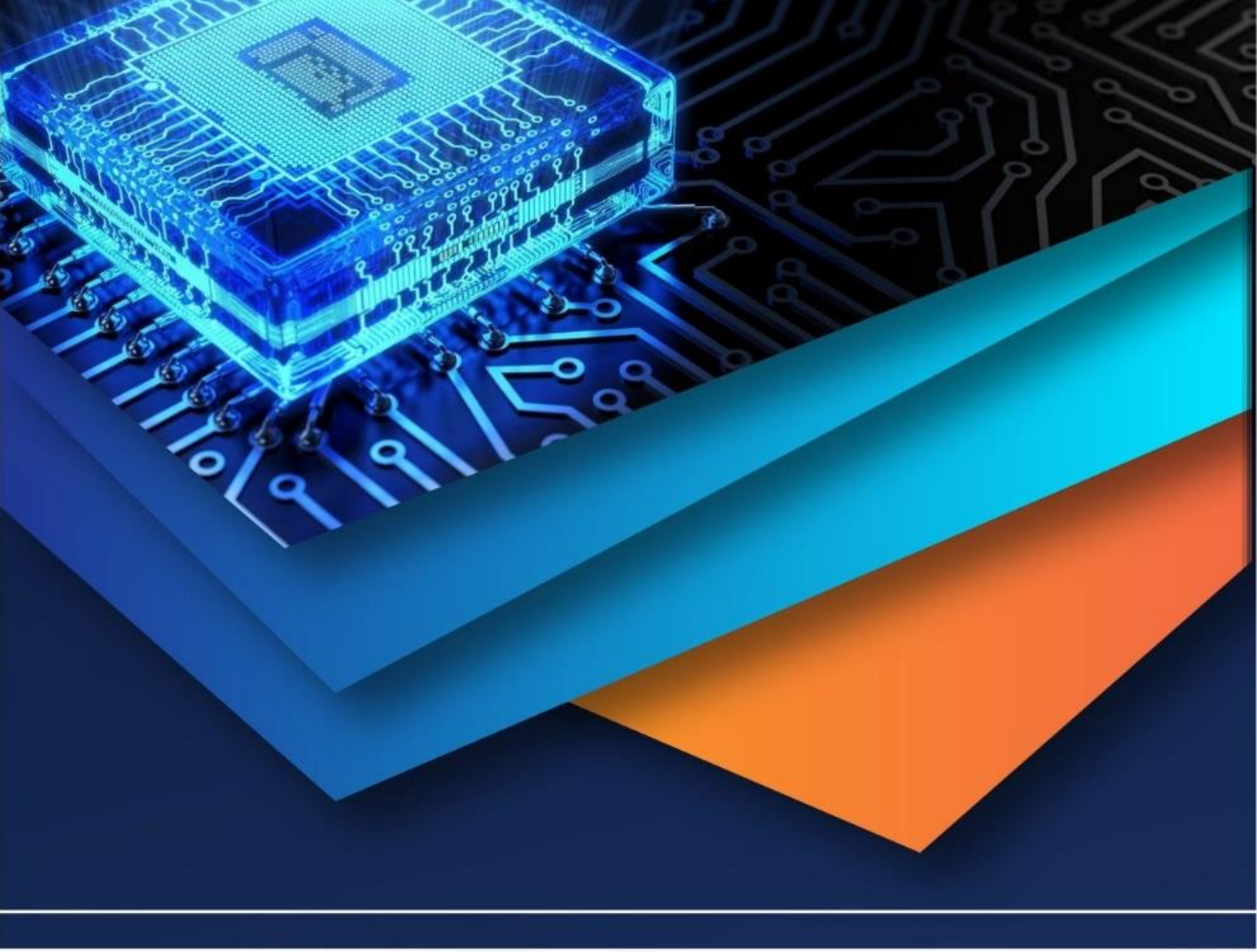
Fig.2.Proposed System Architecture

VI. CONCLUSION

In conclusion, the outlined methodology for a music recommendation system based on facial expressions provides a robust foundation for future implementation. The comprehensive approach covers data acquisition, preprocessing, and model development, leveraging state-of-the-art techniques in deep learning, activation functions, and model optimization. The proposed integration of React.js, React Native, TensorFlow Serving, and deployment on GCP underscores a commitment to delivering a scalable and user-friendly experience. While the system remains in the conceptual stage, the strategic gathering of information and methodologies positions us well for the subsequent stages of implementation. Future endeavours will focus on translating these insights into a functional and innovative music recommendation system that effectively blends emotion detection with personalized playlist suggestions.

REFERENCES

[1] AHLAM ALRIHAILI, ALAA ALSAEDI, KHOLOOD ALBALAWI, LIYAKATHUNISA SYED; “Music Recommender System for Users Based on Emotion Detection through Facial Features”; 2019 12th International Conference on Developments in eSystems Engineering (DeSE).
[2] B. C. KO, "A Brief Review of Facial Emotion Recognition Based on Visual Information", sensors, vol. 18, pp. 401, 2018.
[3] P. VIOLA AND M. JONES, "Rapid object detection using a boosted cascade of simple features", Computer Vision and Pattern Recognition 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, pp. I-I, 2001.
[4] M. A. TURK AND A. P. PENTLAND, "Face recognition using eigenfaces", Computer Vision and Pattern Recognition 1991. Proceedings CVPR’91. IEEE Computer Society Conference on, pp. 586-591, 1991.
[5] NAZIBA MOSTAFA, YAN WAN, UNNAYAN AMITABH, PASCALE FUNG;” A Machine Learning based Music Retrieval and Recommendation System”, Human Language Technology Center Department of Electronic and Computer Engineering.
[6] VINCENZO MOSCATO, ANTONIO PICARIELLO, GIANCARLO SPERLII, “An Emotional Recommender System for Music “University of Naples “Federico II,” 80125 Naples, Italy
[7] SARTHAK SENGUPTA, DR. ANURADHA KONIDENA, “Generating Music Therapy Using Deep Learning and Reinforcement Learning “International Journal of Engineering Applied Sciences and Technology, 2020 Vol. 4, Issue 12, ISSN No. 2455-2143.
[8] K DEVENDRAN, S K THANGARASU, P KEERTHIKA, R MANJULA DEVI, B K PONNARASEE, “EFFECTIVE PREDICTION ON MUSIC THERAPY USING HYBRID SVM-ANN APPROACH” ITM Web of Conferences 37, 01014 (2021).
[9] PRANJUL AGRAHARI AYUSH SINGH TANWAR, BIJU DAS, PROF. PANKAJ KUNEKAR,” Musical Therapy using Facial Expressions” International Research Journal of Engineering and Technology (IRJET) Volume: 07 Issue: 01 | Jan 2020
[10] HORIA ALEXANDRU MODRAN, TINASHE CHAMUNORWA, DORU URSUTIU, CORNEL SAMOILĂ, HORIA HEDESIU, “Using Deep Learning to Recognize Therapeutic Effects of Music Based on Emotions”, Sensors 2023



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)