International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Sounds: Orchestrating Emotions with Machine Learning for Personalized Song Recommendations

Kalyani Kadam[1], Rutuja Kadam[2], Pooja Bagane[3], Akshansh Sourabh[4], Saakshi Jha[5], Srushti Nagrale[6], Saumya Pandey[7]

Submitted: 27/11/2023 **Revised**: 07/01/2024 **Accepted**: 17/01/2024

Abstract*:* Music has a major impact on emotions and mood management, and it frequently helps people cope with the obstacles they face on a daily basis. Finding music that perfectly captures one's present mood might be difficult, though. Currently available music recommendation systems generally rely on user preferences and listening history, which may not be useful for users looking for music to match their current feelings. This study suggests a brand-new method of providing personalised song recommendations by analysing facial expressions to ascertain the user's current mood. The system seeks to give a more individualised and organic way to finding music that corresponds with the user's emotions by utilising computer vision and machine learning techniques. Furthermore, as music may be a useful tool for self-soothing and emotional control, this technique may be especially helpful for people who are struggling with mental health concerns. In order to provide a reliable and robust music listening experience, the efficacy and accuracy of the proposed system will be thoroughly tested and user feedback on a large dataset of facial expressions and song suggestions will be collected.

Keywords: *Emotional well-being, Facial expression analysis, Machine learning, Mood regulation, Music recommendation, Personalized music.*

1. Introduction

Music has the remarkable ability to evoke emotions and influence our mood. When faced with difficulties in daily life, many people use music as a coping method to find comfort and emotional support. However, it might be challenging to choose the ideal music that matches one's present feelings. [1] The majority of music recommendation algorithms now in use are based on past listening patterns or user preferences, however they may not be useful for those looking for music that matches their current mood. We suggest an emotion-based music suggestion system that operate on facial behavior analysis to ascertain the user's current emotional state in order to alleviate this issue. Essential emotional indicators may be immediately derived from the user's facial expressions, allowing the system to reliably estimate the user's mood. The time-consuming and difficult work of manually classifying music into various emotional categories is subsequently relieved by the algorithm, which creates a customized playlist of songs that are in line with the identified emotional state.

We can properly determine the user's current mood by extracting key information from their facial expressions. The implementation of this emotion analysis in music recommendation allows for the curation of a playlist of songs in tendering with the emotional behavior of the user.

[1,3,4,5,6] *Department of Computer Science, Symbiosis Institute of Technology, (SIT) affiliated to Symbiosis International (Deemed University), Pune, India ORCID ID : 0000-0001-9611-9601*
[6] *Vishwakarma Institute of Information Technology, Pune, India*
** Corresponding Author Email: poojabagane@gmail.com*

[2] By automating the laborious and time-consuming process of manually classifying music into different emotional themes, this technique enables the production of personalised playlists that connect with the person's emotional characteristics. This system's technology analyses the user's facial behavior to discover or determine their current mental state and creates a playlist that corresponds to those feelings.

1.1. Project Idea & Proposition

Mainstream music predominates in recommendations in the age of social media platforms like TikTok and Instagram, resulting in fewer customised music experiences on streaming services. A recommendation system for music that is based on emotions is also necessary since consumers frequently utilise music to express their emotions. Existing methods employ audio signals and cooperative filtering to make music recommendations based on a user's prior listening preferences. Through an interaction with a chatbot, this research study suggests a personalised system that assesses the user's current emotions. The chatbot gauges the user's mood by asking a number of open-ended questions, and a corresponding score is computed using the user's input. The technology then uses computer vision and machine learning to analyse facial expressions in order to authenticate the user's mood and create a playlist that appropriately captures their emotional state.

1.2. Motivation

Numerous studies have shown that music has a substantial influence on our emotions and brain activity. Music is an

essential component of human life. Many people use music to manage a variety of emotional states. The ambition to improve the music listening experience by offering personalised playlists that take into account the user's emotions is the driving force behind this initiative. Users may find it difficult to manually create and organise playlists, particularly at certain emotionally charged occasions.

We want to provide individualised music recommendations that take into account each user's unique emotional requirements and musical tastes by creating an emotion-based song recommendation system. The potential advantages of this method include enhancing music discovery as well as promoting emotional balance and mental wellness, as music is a great medium for regulating and expressing emotions. your paper.

2. Literature Review

The SVM Algorithm was utilized by Ambara G, Sahil Choudhury, Krupa K S and Kartikey Rai, to develop their Emotion Based Music Recommendation System [1]. The primary goal is to ascertain the user's emotional condition and play the appropriate song in response. The user's facial expressions are captured using a camera, and the emotion is compared to a predefined collection of image datasets that includes all the emotions.

The goal of Amaan Shaikh, Bhushan Patil, and Tejas Sonawane's project is to build a chatbot song recommender system based on machine learning [2]. It is a comprehensive chatbot service where the user can engage in informal chats with the chatbot and receive music suggestion based on the user's voice. The Last.fm API and IBM Tone Analyzer API are both used in the music recommendation feature. Additionally, it makes advantage of the numerous Python libraries that are accessible to create a productive system.

The application of Natural Language Processing, SENN, and CNN to a "Emotional Tone Analysis to Recommend Songs Using a Chatbot " is attempted to be explained by Sudarsanam and Ayaan Gouse [3]. They have stressed how customer experience is improved by interactions between people and services (chatbot applications). They give businesses new strategies to increase customer engagement and operational effectiveness while lowering traditional customer service costs.

J. James Anto Arnold, H. Immanuel James, J. Maria Masilla Ruban, R. Saranya and M. Tamilarasan, investigate how essential the human face is in determining a person's mood. [4]This feed in is highly to be utilize along with other raw inputs, to dig out the data that can be further used to determine a person's mood. Songs are chosen based on the emotion deduced from the gathered input. The goal of the Facial Behavior Based Song Listener is to survey and appraise the information before making a playlist based on that criteria. Their suggested system looks on identifying human behavior in order to design song players based on those emotions. It discusses the methods currently utilize by song players to identify emotions, the methodology our song player uses to identify human behavior, and why it is preferable to utilize our software system for behavior detection.

'Automatic Facial Expression Recognition System' was proposed by Anagha S. [5]. This system comprises three phases which consist of Face recognition then lies the Expression recognition which is followed by feature extraction. Using an RGB colour model, lighting adjustments to acquire the face, and structural algorithms to maintain the important distinguishable facial characteristics, such as the lips and eyes, face identification is performed in the first step. This method also uses the Active Appearance Model Method (AAM) to extract face traits. This method involves locating the points on the face, such as the eyes, brows, and lips, and creating a data file with details on the model points that were found. In order to establish how the AAM Model should evolve, the approach also recognises faces and utilises an expression as input.

Woori Han, Youngseop Kim and Yong-Hwan Lee suggested a technique emanate from Bezier curve fitting [6]. This approach employs a dual step procedure to first recognise and evaluate the face region in the original picture as input, and then it validates the facial expression of certain characteristics in the region of interest [5]. Following the breginning step of facial behavior recognition, that make the use of colour still pictures based on skin colour pixels and initialised three-dimensional filtering, feature maps were used to appraise face position and the placement of the eyes, nose, mouth on the face. In direction to apply a Bezier curve to the facial component, this method technique first bring out the region of interest before extracting information from the feature map. In direction to determine the Hausdorff distance with the help of a Bezier curve between the input facial data and the database picture, this technique uses training.

Jukka Holm, Arto Lehtiniemi, and others proposed a system for dynamic mood images in music suggestions [7]. Users can receive music suggestions depending on the genre and feature of each pictures by using a sets of photos which is present in the database. This system for suggesting music was developed by the Nokia Research Centre. This method uses textual meta tags and audio signal processing to characterise the genre.

Fully automatic facial expression and identification was proposed by A. Pruski, C. Maaoui, and F. Abdat using a three-step face recognition, facial behavior extraction, and facial behavior classification [8] approach. To recognize the facial feature information, this approach append the Shi and

Thomasi technique with an anthropometric model. In this approach, the classification is based on SVMs i.e, Support Vector Machines, which describe facial traits from neutral faces using a range of 21 distances.

Fang-Fei Kuo and Suh-Yin Lee. [9] Users can gain from the growth of song suggestions as digital music becomes more prevalent. The current suggestion techniques are based on the consumers' musical interests. But there are times when choosing song based on behavior is essential. In this study, we provide a novel method for association discovery from music recommendations based on emotional film scores. We investigated the extraction of musical features and controlled the affinity graph to identify correlations between feelings and musical aspects. Investigational results show that the anticipated method has an average accuracy of 85%.

A substantial amount of research in recommender systems, according to Ivana Andjelkovic and John O'Donovan [10], efforts on developing expectation and ranking. The significance of other traits of the implications, such as transparency, control, and overall user experience which however, been highlighted by recent research. MoodPlay is a hybrid song suggestion system which make use of content and feeling based sorting in a fun user interface based on these qualities. We reflect users that how MoodPlay is used to browse song collections by emotional behavior and we go through how to mix user input at suggestion time with likelihoods based on a previous user profile. Review of the results of a user research (N=240) that examined four situations with various levels of visualisation, interaction, and control.

Anukriti Dureha [11] suggested manual playlist segmentation and song annotation, which can be laborious and time-consuming depending on the user's emotional state at the time. A multitude of techniques have been developed to automate this operation. Nevertheless, the existing algorithms are less precise, less efficient, and need more hardware, which drives up the overall cost of the system. This article offers an algorithm that use a user's facial expressions to create a music playlist automatically, saving time and work compared to a human process. The method proposed in this work attempts to reduce the overall cost and computation time of the system. It also aims to increase the correctness of the suggested method. The facial expression recognition module of the suggested approach is evaluated against user-dependent and user-independent datasets to ensure its accuracy.

Bruce Ferwerda and Markus Schedl [12] make the use of the suggestions that including personality and emotional behavior into songs selections would enhance the fundamental research hypothesis. It is believed that by accounting for these psychological components, the accuracy of the suggestions may be improved. The technique put efforts on the relation between an individual feelings and their use of song as a system for emotional directive.

According to Erdenebileg Batbaatar, Keun Ho Ryu and Meijing Li [13], emotion detection will play a promising role in the field of artificial intelligence and human-computer interaction. From an application perspective, computational linguistics is becoming more and more dependent on the capacity to recognise human emotions in a text. A brand-new dual neural network model is put out in this paper. They created a semantic and emotion sentence encoder using BiLSTM and CNN, respectively. CNN is made to successfully extract emotion information, while BiLSTM is made to gather contextual information. They employed the fine-tuning technique on pre-trained word embeddings to gain a deeper understanding of the semantic and emotional information on a particular dataset, which effectively enhances the performance of emotion detection models from text. In order to extract emotion from the text, we next concatenated the sentence-level encoded vectors.

A novel method for playing music automatically while using facial emotion has been proposed by Madhuri Athavle, Deepali Mudale, Upasana Shrivatav, and Megha Gupta [14]. According to them, the bulk of existing approaches include manually playing music, donning wearable computers, or classifying according to auditory characteristics. Instead, we advise switching the manual playing and sorting. For emotion recognition, they used a convolutional neural network. Pygame & Tkinter are used to suggest music. Their suggested system has the potential to decrease the system's overall cost and processing time requirements, improving the system's overall accuracy.

Researchers Ahmed Al Marouf, Md. Rahmatul Kabir Rasel Sarker, Rafayet Hossain, Bishwajeet Pandey, and Shah Md. Tanvir Siddiquee [15] investigated the feasibility of utilising IBM Watson Tone Analyser, an API service, to extract linguistic and emotional tones from song lyrics. They used a machine learning approach to extract the characteristics from a dataset of 300 English songs and categorise the verbal tone and emotional tone. For classification, they employed several classifiers, such as Naive Bayes, decision trees, random forests, sequential minimum optimisation, and basic logistic regression.

The research work proposed by Nair A., Pillai S., and Nair G.S.[16] creates a personalized system in which the chatbot is used to analyze the user's present emotion. Through the use of several general inquiries, the chatbot determines the user's mood. A score is generated for each response based on the user's input, adding up to a total score that is utilized to create the playlist. The Spotify platform and API are used by the suggested recommendation system to create and recommend playlists.

This study by Jaison C., Raichal Rose Rajeswari, M.[17] presents a song suggestions technique that uses artificial intelligence and machine learning to provide customised music recommendations based on the user's emotional state. The system uses MFCC to extract characteristics in order to analyse the user's tone. The recommended method trains the model more accurately by using deep learning models, such as Artificial Neural Networks. In order to create the information needed for the recommendation process, it is crucial to accurately and dependably recognise the user's emotional state from speech. The recommended technique offers up new avenues for study in the fields of artificial intelligence and music recommendation, with the potential to alter how people listen to music.

This research paper by Vayadande, Kuldeep Narkhede, Parth Nikam, Srushti Punde, Nikita Hukare, Sejal Thakur and Rohit [18] proposes a method that make the use of computer vision techniques to appraise a person's facial expressions and recommends songs based on the behavior that are expressed in the picture. The device make use of integrated camera to capture the user's facial behavior. A machine learning algorithm that has been trained on a dataset of these facial behavior and the facial expressions that go along with them is used to process the captured information. The algorithm's objective is to read a person's facial expressions to determine their emotional state, then choose music that fits that mood. The system has an accuracy rate of 84.82% when it comes to determining an individual's emotional state and making music recommendations based on their facial expressions. It is a trustworthy and effective tool for making music recommendations. Since the camera is already built into the system, it eliminates the need for additional hardware for facial expression capture, making the system convenient and simple to use. The proposed system has the potential to be used in a wide variety of applications, including song therapy, personalized song recommendations for specific individuals, and even in public settings like malls or airports where music can be used to lift spirits of those present. Overall, the suggested system is a novel method of personalized music recommendations that considers a person's emotional state and enables them to find new music that matches their current mood. Integrate physiological signals to boost music recommendation's precision and efficiency.

3. Implementation

In this research, our first step is to gather a diverse collection of face photos showcasing various emotions. To ensure a reliable dataset, we will explore publicly available sources or employ crowd-sourcing platforms, all with the consent of participants. These images will undergo preprocessing to maintain consistency and enhance training accuracy. The preprocessing steps will involve face detection and alignment, resizing the images to a standard resolution, and normalizing them to a common color space, providing a suitable input for the subsequent deep learning model.

For the development of the emotion recognition model, we will utilize a deep learning framework, such as TensorFlow or PyTorch. The model will be designed to identify facial emotions from the preprocessed images. To effectively learn and recognize facial features and emotional patterns, the model architecture will be based on Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs). We will train the model using the collected dataset, employing appropriate loss functions and optimization algorithms to achieve optimal performance and accurate emotion recognition.

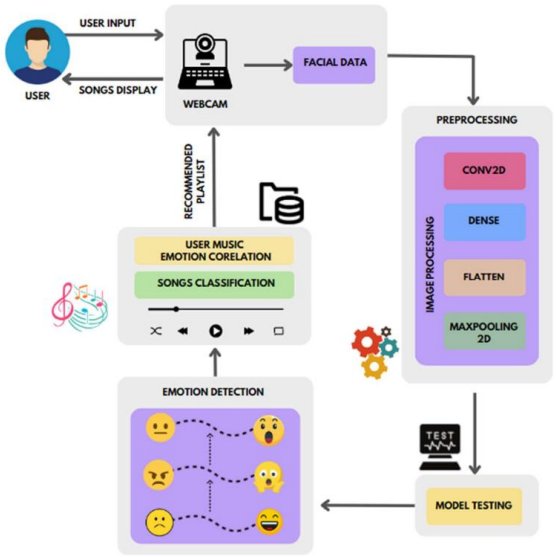

Fig 1. Block diagram of the system design

To systematize the process of recommending music based on users' emotional states, we will create a comprehensive database of songs and their corresponding emotional values. These values can be manually annotated or sourced from existing music databases that provide metadata related to emotions evoked by each song. Each song in the database will be associated with emotional labels, allowing the music recommendation system to make accurate selections based on users' emotions.

For real-time music recommendation based on users' emotional states, we will seamlessly integrate the trained emotion recognition model with the music recommendation system. When a user interacts with the system, their facial expression will be captured using a camera interface. The emotion recognition model will analyze the image and predict the user's current emotional state. Based on this emotional state, the music recommendation system will retrieve songs with corresponding emotional labels, offering personalized music suggestions tailored to the user's expressed emotions.

To provide a user-friendly experience, we will create a web-based user interface using a framework like Flask. The interface will display the suggested music recommendations based on the user's current emotional state, accessible on the user's device. The web framework will enable real-time facial expression capture through the integrated camera, allowing for instant emotion analysis and accurate music recommendations.

To evaluate the effectiveness of the developed system, we will conduct testing with a sample group of users. Participants will interact with the system and provide feedback through surveys and assessments. We will analyze the collected data, including metrics like recommendation accuracy, user satisfaction, and response time. Based on this feedback and evaluation, we will refine and improve the system to enhance its performance and overall user experience.

In conclusion, our implementation plan outlines the development of an emotion-based music recommendation system. By integrating emotion recognition capabilities with a comprehensive music database and user-friendly interface, the system aims to provide personalized music suggestions based on users' emotional states, creating a tailored and enjoyable musical experience for each user.

4. User Interface

In the first phase of our project, we made a Graphical User Interface for camera feature that records the user's expression and feeds the data to the trained model so it can ascertain the user's emotional condition.

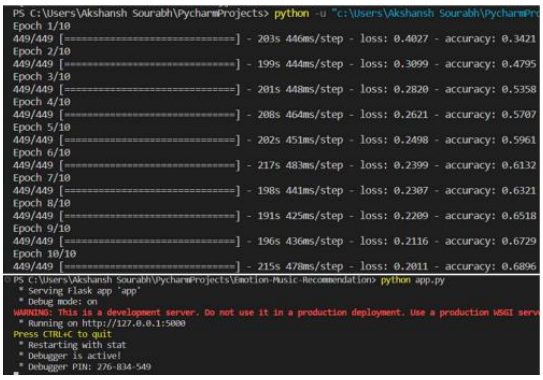

Fig 2. GUI showing the accuracy of our model

When the system is deployed on a web server using flask it will show such an interface to the user. A camera dialogue is there for scanning the user's facial data using computer vision and then it generates the required result based on our machine learning algorithm.

Fig 3. GUI showing the final output of the application

5. Results & Discussion

The goal of the proposed research project is to create an emotion-based song recommendation system that uses facial expression data to assess the user's current emotion and make tailored music recommendations. The system uses a neural network model with a number of layers, including dense, flatten, sequential, conv2d, maxpooling2d, and Adam optimizer, to accurately detect the user's emotional state. Real-time facial expressions of the user are captured by the system's built-in camera and put into a trained model, which produces accurate emotion identification results.

6. Conclusion

In this work, we suggested a model for song suggestions based on emotional facial behavior. This study presented an emotion-based facial recognition-based music recommendation system. Any tension or emotion can be relieved by listening to music. The possibility for constructing suggestion systems for songs based on emotions behavior has recently increased. Therefore, the suggested system offers a face-based emotion detection system to identify emotions and play music depending on those emotions. Increased efficiency can be achieved by using more facial data and train the model again. In the same system, Text & Speech vice versa to our system for more efficient and transparent UI. Our accomplishment proves that this system is a viable alternative to traditional recommendation systems. Indeed, we believe that advances in task-oriented system have enormous potential to improve customer experience and drive business growth through previously untapped channels.

7. Future Work

The suggested emotion-based music recommendation system, which employs machine learning and facial expression analysis, lays the groundwork for future improvements to its features and usability. The possibility for future growth can be seen in the following areas:

1. Database Integration: The system may be coupled with a database to store and manage song data, increasing scalability and efficiency. This would allow for smooth

integration of new music into the recommendation pool and real-time changes.

2. Periodic Playlist Refresh: Adding a tool to automatically update playlists with more recent and better recommendations would guarantee that customers always get the most recent music recommendations.

3. Seamless Music Playback: A feature that allows instant playback or direct referral to suggested tracks on well-known music streaming services like Spotify can be enabled.

4. Client-Side Video Streaming: By making the code more responsive and using less server-side processing, client-side video streaming will improve system responsiveness.

5. Web-Based Application Deployment: The system will be more accessible to a wider audience if it is installed on public servers as a web-based application.

6. User Profiles and customisation: Including user profiles and customisation capabilities can increase the accuracy of recommendations based on unique listening preferences.

The emotion-based music recommendation system may develop into a more sophisticated and user-centric tool by investigating these possible improvements, offering a richer and more individualised music discovery experience.

References

[1] Ambara, Kartikey Rai and Sahil Choudhury, "Implementation of Emotion based Music Recommendation System using SVM Algorithm", 2020, e-ISSN: 2395-0056, p-ISSN: 2395-0072

[2] "Chatbot Song Recommender System", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.8, Issue 12, page no.d222-d226, December-2021

[3] Mr. Sudarsanam and Ayaan Gouse, "Emotional Tone Analysis to Recommend Songs using a Chatbot, ISSN (Online) 2581-9429, Vol. 2, Issue 6, May 2022

[4] H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tamilarasan, R. Saranya, "EMOTION BASED MUSIC RECOMMENDATION SYSTEM", e-ISSN: 2395-0056, p-ISSN: 2395-0072, Volume 06, Issue 03, Mar 2019

[5] Anagha S.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" 2014 International Conference on Electronics and Communication System (ICECS -2014).

[6] Yong-Hwan Lee , Woori Han and Youngseop Kim, "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting" 2013 16th International Conference on Network-Based Information Systems.

[7] Arto Lehtiniemi and Jukka Holm, "Using Animated Mood Pictures in Music Recommendation", 2012 16th International Conference on Information Visualisation.

[8] F. Abdat, C. Maaoui and A. Pruski, "Humancomputer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer

[9] T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation," J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[10] Renuka R. Londhe, Dr. Vrushshen P. Pawar, "Analysis of Facial Expression and Recognition Based On Statistical Approach", International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[11] Anukriti Dureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions" : IJCA 2014.

[12] Bruce Ferwerda and Markus Schedl "Enhancing Music Recommender Systems with Personality Information and Emotional States": A Proposal: 2014.

[13] Erdenebileg Batbaatar, Meijing Li, Keun Ho Ryu, " Semantic-Emotion Neural Network for Emotion Recognition From Text", IEEE Access, vol. 7, pp. 111866 - 111878, 2019.

[14] Athavle, Madhuri. (2021). Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE). 2. 1-11. 10.54060/JIEEE/002.02.018.

[15] Al Marouf, R. Hossain, M. R. Kabir Rasel Sarker, B. Pandey and S. M. Tanvir Siddiquee, "Recognizing Language and Emotional Tone from Music Lyrics using IBM Watson Tone Analyzer," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 2019, pp. 1-6, doi: 10.1109/ICECCT.2019.8869008.

[16] Nair, S. Pillai, G. S. Nair and A. T, "Emotion Based Music Playlist Recommendation System using Interactive Chatbot," 2021 6th International Conference on Communication and Electronics Systems (ICCES), Coimbatre, India, 2021, pp. 1767-1772, doi: 10.1109/ICCES51350.2021.9489138.

[17] JAISON C, R. R.; RAJESWARI, M. Song Recommendation based on Voice Tone Analysis. 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Electronics and Renewable Systems (ICEARS), 2023 Second International Conference on, [s. l.], p. 708–712, 2023. DOI 10.1109/ICEARS56392.2023.10085001. Disponível em: https://search.ebscohost.com/login.aspx?direct=true&

AuthType=ip,sso&db=edseee&AN=edseee.10085001 &site=eds-live. Acesso em: 4 ago. 2023.

[18] VAYADANDE, K. et al. Facial Emotion Based Song Recommendation System. 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), Computational Intelligence and Sustainable Engineering Solutions (CISES), 2023 International Conference on, [s. l.], p. 240–248, 2023. DOI 10.1109/CISES58720.2023.10183606. Disponível em: https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,sso&db=edseee&AN=edseee.10183606 &site=eds-live. Acesso em: 4 ago. 2023.