IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 6, June 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

Musical Moods: Emotion Detection and Music Recommendation

Sameera Fathima, Dr. C D Guruprakash

2nd Year, Master of Technology, Department of Computer Science Engineering, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India

Professor and HOD, Department of Computer Science Engineering, Sri Siddhartha Institute of Technology, Tumakuru, Karnataka, India

ABSTRACT: Recommender systems help users find the best possible options for them by guiding them through a large space of possible objects. The recommender systems need to accurately capture customer needs and preferences in order to provide reliable recommendations. But for subjective and complex products such as musics, music, and news, user emotions play an unexpectedly critical role in the decision process. Our new Emotion-based Music Recommender System can better understand and predict the changing preferences of users, making it a better solution for recommendations. The system's objective is to offer consumers' individualized recommendations based on their interests and preferences. The suggestion is predicated on the notion that the user is experiencing a particular emotion and favors a particular kind of material as well as feedback from comparable users. An effective music recommendation system is presented, which recommends musics based on the real-time mood of the user. It mainly focuses on a Convolutional Neural Network (CNN) model based on the Mobile Net architecture, capable of classifying 7 different human facial emotions. The system consists of three modules: the Emotion Module, Music Classification Module, and Recommendation Module. The Emotion Module takes an image of the user's face as input and utilizes CNN to identify their present mood. The Music Classification Module categorizes musics into genres based on their content features. The Recommendation Module suggests personalized music recommendations to the user based on their real-time mood and music

KEYWORDS: EmotionDetection,MusicRecommendation,SQLite(StructuredQueryLanguageLite),Librosa,CNN(Convolutional Neural Network),Wed Scraping,Flask,Analysis

I.INTRODUCTION

Recommendation system is essentially a kind of data filtering. Recommendation systems are very important in today's world since they give users suggestions based on their interests and needs when they are looking for something. Music recommendation systems aim to assist music lovers by recommending what music to watch without requiring them to go through the time-consuming and complex process of selecting from a vast number of musics ranging from thousands to millions. One of the most important social cues is a person's face. Particularly. facial expressions convey the emotion that a target is now feeling, which in turn affects how the target is seen and what behavioural inclinations the observer is motivated to exhibit. Hence extracting and appreciation of emotion has high importance on the interaction between human and the computing device communication. The number of facial behaviours and motions can be parameterized primarily based on muscle actions. AI can recognize emotions by learning and understanding what each facial expression means and applying that knowledge on the new information presented to it. Humans can recognize and distinguish between faces. Now computers are able to do the same. This opens heaps of applications.

The project develops a user-centric music recommendation system using advanced technologies like facial emotion recognition, voice pitch analysis, and user input. Utilizing CNN mobile architecture and Librosa for accurate recommendations, the system offers a sophisticated interface for discovering and enjoying music tailored to users' emotions and preferences.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

Applications

1. **Personalized Music Recommendations:** The primary application of the model is to provide personalized music recommendations based on user emotions.
2. **Enhanced User Engagement**: Incorporating facial emotion recognition into the music recommendation process can significantly enhance user engagement. By allowing users to input their emotions directly through facial expressions, voice, or text, the system provides an interactive and immersive experience.
3. **Emotion-Based Content Filtering:** The incorporation of emotion recognition technology can also be extended beyond music recommendations. The same underlying concept can be applied to filter and categorize musics based on their emotional content.
4. **Market Research and Audience Insights:** The data collected through the music recommendation system presents an opportunity for market research and audience insights.

Research motivation and Problem Statement

This research aims to develop user-centric and emotionally-driven music recommendation systems by integrating facial emotion recognition, voice pitch analysis, and text-based input. Emotions significantly shape individuals' music preferences and experiences, and emotion recognition technology has potential applications in various domains. The research aims to enhance the music experience by providing personalized, emotionally-aligned recommendations that resonate with users' current emotional state. By incorporating multi-modal inputs and leveraging emotion recognition techniques, this research contributes to the development of sophisticated and user-centric recommendation systems.

Traditional music recommendation systems lack the ability to capture and incorporate user emotions effectively, relying solely on explicit feedback. Current emotion recognition technologies often focus solely on facial expressions, neglecting additional cues from voice pitch and textual input. This project aims to address these limitations by developing a music recommendation model that utilizes facial emotion recognition, voice pitch analysis, and text-based input. The goal is to enhance the music-watching experience by providing personalized recommendations aligned with the user's emotional state.

Research Objectives

1. **Develop a Multi-Modal Input System**: Create a robust and user-friendly system that can accept input through multiple modalities, including facial expressions, voice pitch, and text. To Help music service providers maintain their client base and attract new users by boosting customer happiness.
2. **Build an Emotion-Aware Music Recommendation Model**: Design and implement a music recommendation model that integrates the multi-modal user inputs with a comprehensive dataset.
3. **Evaluate and Refine the Recommendation Model**: Conduct extensive evaluations and performance testing of the recommendation model to measure its accuracy, effectiveness, and user satisfaction. Incorporate feedback and iterate on the model to enhance its recommendation accuracy and overall user experience.
4. **Implement a User-Friendly Interface**: Create an intuitive and user-friendly interface that allows users to input their emotions conveniently, whether through facial expression analysis,voicepitch analysis, or text-based input.

II. LITERATURE SURVEY

Dhavalikar et al., [1] The proposed Automatic Facial Expression Recognition System (AFERS) consists of three stages: face detection, feature extraction, and facial expression recognition. Face detection uses YCbCr color model, AAM method for facial features, and Euclidean Distance method for recognition. This system is crucial in interpersonal relations, revealing affective, cognitive, personality, intention, and psychological states..

D. Garg et al., [4] The paper presents a music recommendation system using collaborative filtering and Apache Mahout. A hybrid filtering strategy could enhance the system, providing more accurate and effective recommendations

Y.Li et al., [5] This paper introduces a hierarchical architecture using a dynamic Bayesian network to track face features and recognize facial expressions simultaneously. The model outperformed existing techniques in facial feature

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

tracking and AU detection, but inferred global expression from AU and facial feature point measurements. Despite not recognizing six basic expressions, combining the two approaches improved face feature points and AUs.

A paper by Y.Zhang et al.,[6] investigates the use of dynamic Bayesian networks (DBN) and multisensory information fusion to model and understand temporal behaviors of face expressions in image sequences. An active IR illumination-based system provides accurate visual information under varying lighting conditions and head motion. A dynamic and probabilistic framework, combining DBN with Ekman's facial action coding system (FACS), reduces recognition uncertainty and increases robustness and accuracy. The method reliably and robustly identifies spontaneous facial emotions in various circumstances.

T. Elias et al., [10A study used deep learning to recommend musics based on mood detection. They applied 16 filters and 32 filters, generating 16 feature maps. The input was flattened into a one-dimensional array and fed into a densely connected fully connected layer. The system produced expression classes like surprised, neutral, glad, and angry. The Convolutional Neural Network (CNN) method achieved 95% accuracy, with VGG16 being the top model achieving 99% accuracy.

Human emotions, according to a paper by R.K. Madupu et al[12], Emotion is crucial for good communication, and can take various forms like facial expressions. Non-verbal communication, such as facial expressions, is more effective than spoken communication. In business, recognizing emotions through facial expressions is essential. Image processing techniques, such as SURF feature extraction and convolution neural network, help identify emotions based on facial expressions.

A paper article by Ruchitaa Raj N R et al.,[19] Web scraping is a powerful and efficient method for extracting data from vast amounts of internet content. Python is the most powerful language with built-in libraries and support for third-party open source libraries. Developers can choose a specific library based on their scanning application. This white paper tests various web scraping tools and techniques, analyzing their performance and demonstrating statistical importance. However, data collectors must avoid violating laws that could lead to legal liability. The document covers various aspects of web scraping tools and software, including processing times.

III. IMPLEMENTATION

3.1 Methodology

This application uses emotion recognition to provide personalized music recommendations using facial expression, voice pitch, and text input. It uses a Convolutional Neural Network mobile architecture, web scraping techniques, and Librosa's audio signal processing capabilities. The back end and front end are integrated through Flask, ensuring smooth communication between the user interface and the recommendation model. Music-related data is stored and retrieved using SQLite3. This approach ensures personalized and context-aware music recommendations based on user emotions.he Non-Commercial Licensing and Copyright/License has been thoroughly verified. The datasets are

3.2 Description Process

This section explores the project's description phase, focusing on music recommendations, user preferences, and emotion recognition techniques. It discusses methodologies, data collection, facial emotion recognition, and voice input analysis, and the Flask framework's integration for seamless communication.

3.2.1 Dataset Collection and Pre-processing

The data collection process involved manual gathering of user information, including music preferences and age groups. A Google Form was created to streamline the process, allowing users to provide preferences and age groups. Analysis using statistical techniques like pie charts and bar diagrams provided valuable insights. However, manual data collection was labor-intensive and time-consuming, leading to a strategic transition to web scraping techniques.

3.2.1.1 Transition to Web Scraping: Enhancing Data Collection Efficiency

Web scraping is the process of extracting data from websites using automated tools. Beautiful Soup, a Python library, simplifies the process by parsing HTML and XML documents, enabling the extraction of music genres, ratings, and user reviews. This automated data extraction ensures accurate acquisition for a music recommendation system. The data is stored in gzipped, UTF-8 character-set, tab-separated-values (TSV) formatted files, with headers and

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

placeholders for missing or null fields. This approach improves data collection efficiency, saves time and resources, and ensures a comprehensive and up-to-date dataset for the music recommendation system.

3.2.2 Facial Emotion Recognition

The facial recognition model in this project utilizes CNN mobile architecture and LBP (Local Binary Patterns) for efficient emotion classification. CNN mobile architecture extracts essential features like texture, shape, and patterns, while LBP extracts texture information by comparing pixel intensity values within local neighborhoods. This combination enhances the model's accuracy in detecting and interpreting facial expressions.

3.2.3 Voice Pitch Analysis

The Voice Pitch Analysis section uses Librosa, an audio signal processing library, to analyze voice input. It performs tasks like audio loading, spectrogram computation, and feature extraction. Librosa's key capabilities include mel-spectrograms and MFCCs, which capture frequency content and temporal dynamics. It also extracts emotion-related features, including pitch-related information, for better understanding emotional nuances. This enhances the voice analysis component, capturing acoustic attributes for improved emotion recognition and improving the effectiveness and accuracy of the recommendation model.

3.2.4 Backend and Frontend Integration

The Flask framework is used for backend development and frontend design, providing essential functionalities like routing, request handling, and template rendering. It follows a minimalistic approach, allowing developers to have flexibility and control over the application's structure and components. Flask supports HTTP methods like GET and POST, enabling seamless communication between the backend and frontend. By combining Flask's backend capabilities with HTML and CSS for frontend design, the integration process ensures cohesive functioning, resulting in a user-friendly music recommendation system.

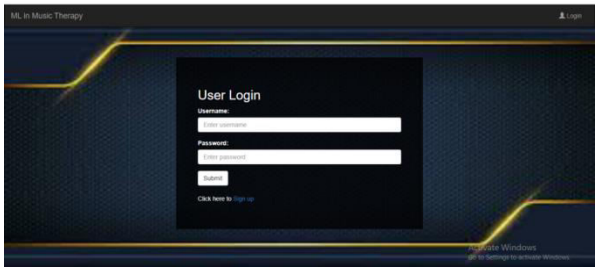
IV. RESULTS

The web application features a login screen with a registration option for new users. After signing in or signing up, users are directed to the main screen where they can choose from three options: facial emotion recognition, voice pitch analysis, or manual input. These choices allow users to input their emotions through facial expressions, voice analysis, or by manually entering their preferences

4.1 User Authentication Screens: Sign In and Sign Up

Figures 4.1.1 and 4.1.2 depict the login page of the web application. Upon successful login or sign up, users are directed to the main screen, where they can access the application's primary features and functionalities.

Fig. 4.1.1 User Sign In Page of the Web Application

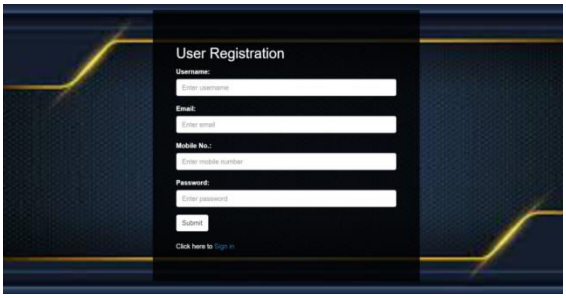


| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

Fig. 4.1.2 User Sign Up Page of the Web Application



4.2 Home Page and Feature Overview

This section provides an overview of the application's main screen, showcasing its intuitive layout and highlighting key features available for users

Fig 4.2.1 Home Screen of the Web Application

4.2.1 presents the home Figure page of the application, featuring three distinct options for user interaction: facial emotion input, voice pitch analysis, and manual text input.

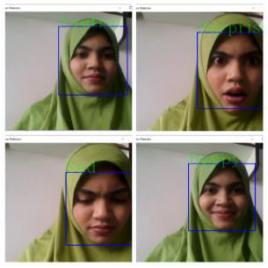


Fig 4.2.2 Emotion Recognizer Classifies Facial Expressions

The depicted figure 4.2.2 demonstrates the emotion classifier effectively distinguishing between neutral,happy, sad, and surprised emotional states, showcasing its ability to accurately identify and classify different emotions based on facial expressions.

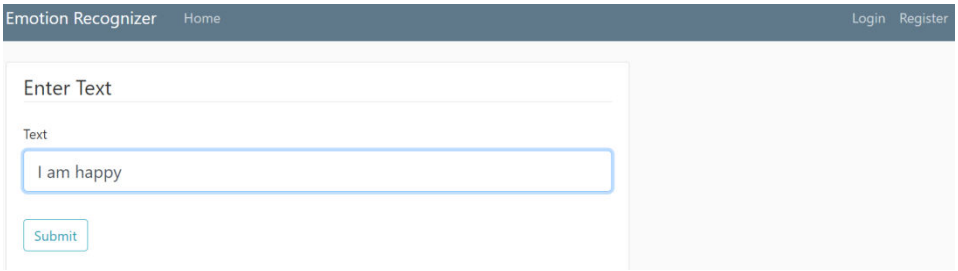


Fig 6.2.3 Happy Emotion input screen through keyboard.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

Fig 6.2.3 depicts the keyboard input screen which allows users to input their emotions manually, providing a convenient and flexible way to express their preferences and emotional states within the application here happy feelings is expressed

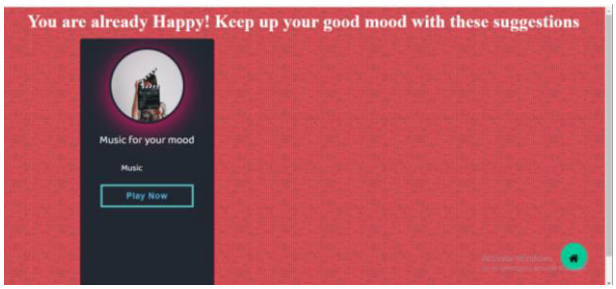


Fig 6.2.4 Landing Page after Successful Recognition of Happy Emotion

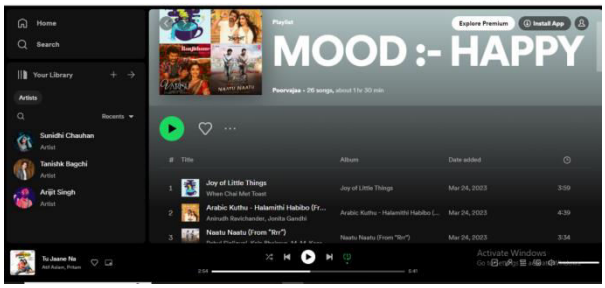


Fig 6.2.5 Music Recommendations for the Emotion 'Happy'

The figures 6.2.4 and 6.2.5 below illustrate the recommendation results specifically generated for the emotion "Happy," showcasing the application's ability to provide personalized music recommendations tailored to the user's emotional state.

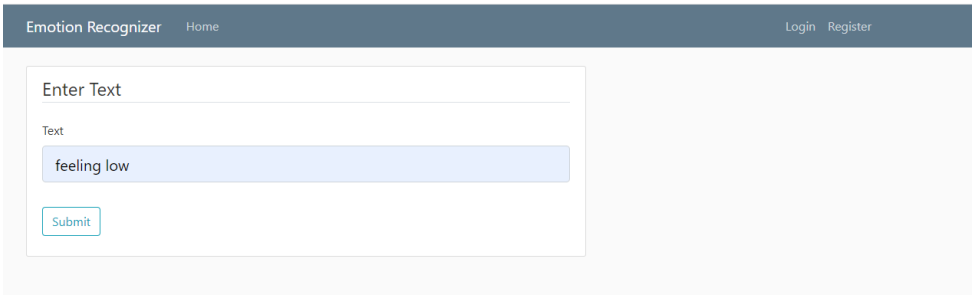


Fig 6.2.6 Emotion input screen through keyboard.

Fig 6.2.6 depicts the keyboard input screen which allows users to input their emotions manually, providing a convenient and flexible way to express their preferences and emotional states within the application here sad feeligs is expressed

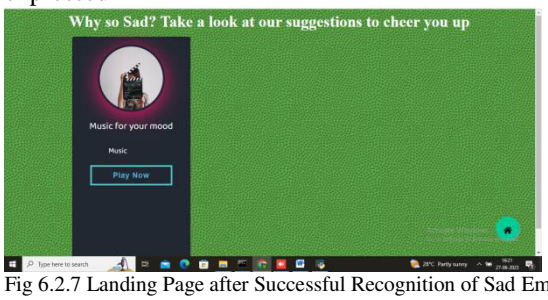


Fig 6.2.7 Landing Page after Successful Recognition of Sad Emotion

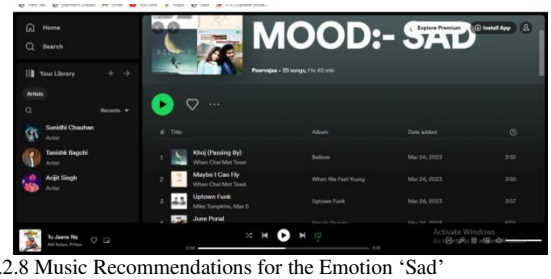


Fig 6.2.8 Music Recommendations for the Emotion 'Sad'

The figures 6.2.7 and 6.2.8 below illustrate the recommendation results specifically generated for the emotion "Sad," showcasing the application's ability to provide personalized music recommendations tailored to the user's emotional state.

V. CONCLUSION AND FUTURE WORK

The project has successfully developed a music recommendation system that incorporates facial emotion recognition, voice pitch analysis, and user input. The utilization of the CNN mobile architecture for emotion recognition, Librosa for audio signal processing, and web scraping techniques for data acquisition has resulted in a personalized music recommendation system. The integration of Flask framework for backend development and SQLite3 database management system has ensured seamless functionality and efficient data storage. Through rigorous testing and

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

evaluation, the system has demonstrated promising results in providing relevant and engaging music recommendations to users.

Moving forward, there are several potential avenues for enhancing and expanding the capabilities of the music recommendation system. Firstly, further research and development can focus on improving the accuracy and robustness of the facial emotion recognition component. This can be achieved by exploring advanced deep learning architectures and incorporating more comprehensive emotion models. Additionally, expanding the recommendation system to incorporate multi-modal inputs, such as combining facial expressions with voice analysis or text sentiment analysis, can provide a richer understanding of user preferences and enhance the accuracy of recommendations. Integration of user feedback mechanisms within the system can enable users to provide explicit feedback on recommended musics, allowing for continuous refinement of the recommendation algorithms and more personalized and adaptive recommendations over time.
By pursuing these future objectives, the music recommendation system can continue to evolve and deliver highly relevant and engaging music recommendations to users, enhancing their music-listenings experience and satisfaction.

REFERENCES

[1] A. S. Dhavalikar and R. K. Kulkarni, "Face detection and facial expression recognition system," *2014 International Conference on Electronics and Communication Systems (ICECS),* 2014.

[2] M. Kim, S. Jeon, H. Shin, W. Choi, H. Chung and Y. Nah, "Movie Recommendation based on User Similarity of Consumption Pattern Change," *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE),* 2019.

[3] M. Xiaoxi, L. Weisi, H. Dongyan, D. Minghui and H. Li, "Facial emotion recognition," *2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP),* 2017.

[4] C.-S. M. Wu, D. Garg and U. Bhandary, "Movie Recommendation System Using Collaborative Filtering," *2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS),* 2019.

[5] Y. Li, S. Wang, Y. Zhao and Q. Ji, "Simultaneous Facial Feature Tracking and Facial Expression Recognition," *IEEE Transactions on Image Processing (Volume: 22, Issue: 7, July 2013),* 2013.

[6] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 27, Issue: 5, May 2005),* 2005.

[8] N. Perveen, S. Gupta and K. Verma, "Facial expression recognition using facial characteristic points and Gini index," *2012 Students Conference on Engineering and Systems,* 2012.

[9] X. Jia and F. Liu, "Research on intelligent recommendation system model supported by data mining and algorithm optimization," *2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT),* 2021.

[10] M. M. Reddy, R. S. Kanmani and B. Surendiran, "Analysis of Movie Recommendation Systems; with and without considering the low rated movies," *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE),* 2020.

[11] T. Elias, U. S. Rahman and K. A. Ahamed, "Movie Recommendation Based on Mood Detection using Deep Learning Approach," *2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT),* 2022.

[12] Y. Hou, "Application of Neural Graph Collaborative Filtering in Movie Recommendation System," *2021 IEEE International Conference on Electronic Technology, Communication and Information (ICETCI),* 2021.

[13] R. K. MADUPU, C. KOTHAPALLI, V. YARRA, S. HARIKA and C. Z. BASHA, "Automatic Human Emotion Recognition System using Facial Expressions with Convolution Neural Network," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA),* 2020.

[14] W. Wang, C. Ye, P. Yang and Z. Miao, "Research on Movie Recommendation Model Based on LSTM and CNN," *2020 5th International Conference on Computational Intelligence and Applications (ICCIA),* 2020.

[15] G. Sunandana, M. Reshma, Y. Pratyusha, M. Kommineni and S. Gogulamudi, "Movie recommendation system using enhanced content-based filtering algorithm based on user demographic data," *2021 6th International Conference on Communication and Electronics Systems (ICCES),* 2021.

[16] N. P. Hegde, C. Shetty, B. Dhananjaya, Deepa, N. Rashmi and H. Sarojadi, "Face and Emotion Recognition in Real Time using Machine Learning," *2022 7th International Conference on Communication and Electronics Systems*

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 6, June 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1106057 |

(ICCES), 2022.

[17] G. S. Kalra, R. S. Kathuria and A. Kumar, "YouTube Video Classification based on Title and Description Text," *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS),* 2019.

[18] A. S. Bale, N. Ghorpade, S. Kamalesh, R. B. S, R. R and R. S, "Web Scraping Approaches and their Performance on Modern Websites," *Third International Conference on Electronics and Sustainable Communication Systems (ICESC 2022),* 2022.

[19] I. Camargo-Henríquez and Y. Núñez-Bernal, "A Web Scraping based approach for data research through social media: An Instagram case," *2022 V Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC),* 2022.

[20] S. K. Malik and S. Rizvi, "Information Extraction Using Web Usage Mining, Web Scrapping and Semantic Annotation," *2011 International Conference on Computational Intelligence and Communication Networks,* 2011.

[21] R. R. N. R, N. R. S and V. M., "Web Scrapping Tools and Techniques: A Brief Survey," *2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT),* 2023.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details