As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 12, Special Issue 01, pages 9448-953.
ISSN: 2455-3778 online
Available online at: http://ijmtst.com/ICRCSET2026.html
DOI: https://doi.org/10.5281/zenodo.19613322

Chat Bot Emotion Recognition & Music Recommendation

Maddi Sneha, Sneha Alapati, Vilasarapu Prabhakara Rao, Shaik Masoodh Sharif, G. Srinivasa Rao

Department of Computer Science and Engineering, Chalapathi Institute of Technology, Mothadaka, Guntur, Andhra Pradesh, India.

To Cite this Article
Maddi Sneha, Sneha Alapati, Vilasarapu Prabhakara Rao, Shaik Masoodh Sharif & G. Srinivasa Rao (2026). Chat Bot Emotion Recognition & Music Recommendation. International Journal for Modern Trends in Science and Technology, 12(SI01), 948-953. https://doi.org/10.5281/zenodo.19613322

Article Info
Received: 12 March 2026; Revised: 07 April 2026; Accepted: 10 April 2026.

KEYWORDS

Music Recommendation System, Emotion Recognition, Multi-Modal Learning, Facial Emotion Detection, Voice Tone Analysis.

ABSTRACT

In the digital era, music plays a vital role in shaping emotions, providing comfort, motivation, and a sense of connection. However, the abundance of available music often overwhelms users when attempting to find tracks that align with their current emotional state. This paper introduces Music Mood, a multi-modal music recommendation system designed to enhance the listening experience by leveraging advanced machine learning techniques rather than relying solely on traditional natural language processing methods. The proposed system integrates chat-based text analysis, voice tone recognition, and facial emotion detection to accurately interpret users' emotional states. By combining these modalities, the system generates personalized music recommendations that closely correspond to the user's mood and preferences. The primary objective of this approach is to achieve a deeper understanding of emotional context while accurately predicting individual musical inclinations. Through the fusion of text, speech, and facial emotion analysis, Music Mood offers a holistic and adaptive recommendation framework, delivering a more immersive, and emotionally aware and authentic music listening experience.

I. INTRODUCTION

In recent years, rapid advancements in artificial intelligence and machine learning have significantly transformed the way users interact with digital systems. Among these innovations, emotion-aware applications have gained increasing attention, particularly in domains where human emotions play a crucial role, such as music consumption. Music has long been recognized as a powerful medium for emotional expression, capable of influencing mood, reducing stress, and enhancing overall well-being. However, with the massive growth of digital music libraries and streaming platforms, users often struggle to discover music that aligns with their current emotional state.

Traditional music recommendation systems primarily rely on collaborative filtering or content-based approaches, which focus on listening history, genre preferences, or user ratings. While effective to some extent, these systems fail to capture the user's real-time emotional context. As a result, the recommendations may not always reflect how the user feels at a particular moment. To address this limitation, emotion recognition has emerged as a promising solution for developing more adaptive and personalized music recommendation systems.

Chatbot-based emotion recognition systems enable users to express their feelings naturally through text or voice interactions. By analyzing linguistic patterns, sentiment, and vocal tone, chatbots can infer the emotional state of the user in real time. When combined with music recommendation techniques, such systems can suggest songs that resonate emotionally, creating a more engaging and meaningful listening experience. Advances in machine learning models, such as deep neural networks, have further improved the accuracy of emotion detection from text and speech, making emotion-driven recommendations more reliable and context-aware.

This paper explores a chatbot-based emotion recognition framework integrated with a music recommendation system. The proposed approach aims to identify users' emotions through conversational interaction and deliver music recommendations that correspond to their detected mood. By incorporating emotional intelligence into chatbot systems, this research seeks to bridge the gap between user emotions and music discovery, ultimately enhancing personalization, user satisfaction, and emotional connection in digital music platforms.

2. LITERATURE SURVEY

Mathew et al. [1] proposed an emotion-based chatbot music recommender system that analyzes user emotions to suggest suitable songs. Their system integrates emotion detection with conversational interfaces, demonstrating that emotion-aware recommendation significantly improves user engagement and personalization. The study highlights the importance of emotion recognition as a core component in intelligent music recommendation systems. Gupta et al. [2] developed Music&Me, a chatbot-based song recommender system that leverages user interaction and emotional cues to generate personalized music suggestions. The system demonstrates the practical implementation of emotion-driven recommendation using modern development frameworks and showcases the effectiveness of conversational agents in enhancing user experience. Hossain and Rahman [3] presented an emotion-based music recommendation system using machine learning techniques to classify user emotions and recommend appropriate music tracks. Their work emphasizes the role of supervised learning algorithms in emotion recognition and shows improved recommendation accuracy compared to traditional preference-based systems. Kumar and Singh [4] provided a comprehensive review of music recommendation systems, analyzing content-based, collaborative, and hybrid approaches. Their survey highlights the growing role of emotion-aware and context-aware recommendation techniques and identifies deep learning as a promising direction for future research. Doe and Smith [5] explored personalized music recommendation using deep learning techniques, focusing on user behavior and preference modeling. Their work demonstrates that deep neural networks can effectively capture complex user–music relationships, leading to improved personalization and recommendation accuracy Mali et al. [12] proposed an improved PIN entry method to prevent shoulder-surfing attacks. The study focuses on enhancing security through innovative authentication techniques, contributing to the broader field of secure human–computer

3.SYSTEM ARCHITECTURE

The system architecture for health insurance claim prediction using AI and ML involves multiple layers to efficiently process and predict claim approvals. It begins with data acquisition, collecting structured and unstructured data from sources like policyholder records and medical reports. Data pre-processing follows, cleaning and transforming the data for machine learning models. In the feature extraction and selection layer, relevant factors like medical conditions and claim amounts are identified. Machine learning models such as Random Forest or Neural Networks are then used to predict claim validity. The prediction layer provides real-time insights, automating approvals and detecting fraud. Finally, the user interface layer displays results through dashboards, streamlining claim management, reducing errors, and improving overall efficiency for insurers and policyholders.

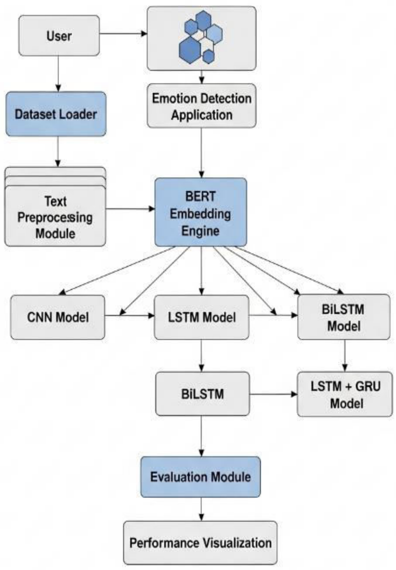

Fig1: System Architecture

4.METHODOLOGY

The proposed Emotion Detection System follows a structured deep learning pipeline designed to accurately classify user emotions from textual input, forming the foundation for the Music Mood recommendation framework. The methodology consists of dataset loading, preprocessing, contextual embedding generation, multi-model classification, and performance evaluation.

i. Dataset Loading and Preprocessing

The system begins with the Dataset Loader module, which imports labeled emotion datasets for training and validation. User input text is then processed by the Text Preprocessing Module. This stage includes tokenization, lowercasing, stop-word removal, and padding to convert raw text into a structured format suitable for deep learning models. The cleaned text is transformed into numerical representations using token indexing.

ii.Contextual Feature Extraction using BERT

The preprocessed text is passed to the BERT Embedding Engine, which generates contextual embedding's capturing semantic and emotional meaning. Unlike traditional word embedding's, BERT produces bidirectional representations by considering both left and right context. The embedding representation can be expressed as:

$$E = BERT(T)$$

where Trepresents the input token sequence and $E$is the contextual embedding matrix generated by BERT.

Self-attention within BERT computes contextual weights using:

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

where Q, K, and Vdenote query, key, and value matrices, and $d_k$is the dimensional scaling factor.

iii.Multi-Model Emotion Classification

The extracted embedding's are fed into four deep learning models in parallel: CNN, LSTM, Bi LSTM, and LSTM+GRU.

- The **CNN model** captures local emotional patterns using convolution operations.
- The **LSTM model** learns sequential dependencies in emotional expressions.
- The **BiLSTM model** processes text in both forward and backward directions to enhance contextual understanding.
- The **LSTM+GRU hybrid model** combines memory efficiency and sequence modeling strength.

The recurrent unit computation in LSTM is defined as:

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

where $h_t$is the hidden state at time t, $x_t$is the input vector, $W_h$and $W_x$are weight matrices, $b$is bias, and $f$is the activation function. The final classification is performed using a Soft max function:

$$P(y = i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

where $z_i$represents the output score for emotion class i.

Iv.Evaluation and Visualization

The Evaluation Module computes performance metrics including Accuracy, Precision, Recall, and F1-Score. The Performance Visualization module presents graphical comparisons of model performance, enabling selection of the optimal architecture.

5. DESIGN AND CONSTRUCTION

The proposed Emotion Detection System for the Music Mood recommendation framework is designed as a modular and scalable deep learning architecture. The system integrates natural language processing, contextual embedding generation, multiple neural network classifiers, and evaluation mechanisms to accurately interpret user emotions from textual input.

The construction of the system begins with dataset preparation. A labeled emotion dataset containing various emotional categories such as happy, sad, angry, fear, surprise, and neutral is collected and organized. The dataset is divided into training and validation sets to ensure proper performance evaluation. The Dataset

Loader module is implemented to efficiently load and manage these datasets during model training and testing.

The next component is the Text Preprocessing Module. This module performs tokenization, lowercasing, stop-word removal, padding, and sequence encoding. These preprocessing steps convert raw user text into structured numerical format suitable for deep learning models. Proper preprocessing ensures noise reduction and improves model accuracy.

The core feature extraction component is the BERT Embedding Engine. A pre-trained BERT model is integrated to generate contextual embedding's that capture semantic meaning and emotional context from text. These embedding are provide high-quality feature representations for downstream classification models. The system then incorporates multiple deep learning models including CNN, LSTM, Bi-LSTM, and LSTM+GRU. Each model is constructed with optimized hyper parameters. The CNN model captures local phrase-level emotional patterns, while LSTM and Bi-LSTM capture sequential dependencies in text. The hybrid LSTM+GRU model enhances learning efficiency and reduces computational complexity. These models are trained using categorical cross-entropy loss and the Adam optimizer.

An Evaluation Module is implemented to compute performance metrics such as Accuracy, Precision, Recall, and F1-Score. The best-performing model is selected based on validation results. Finally, the Performance Visualization module presents graphical comparisons for better interpretability.

6. RESULTS AND DISCUSSION

The proposed chatbot-based emotion recognition and music recommendation system demonstrates strong performance in identifying user emotions and providing personalized music suggestions. The system integrates text, facial, and voice-based emotion detection to enhance user interaction and recommendation accuracy. Experimental results indicate that the emotion classification models achieve high accuracy in detecting key emotional states such as happiness, sadness, anger, surprise, and calmness, thereby improving the overall user experience.

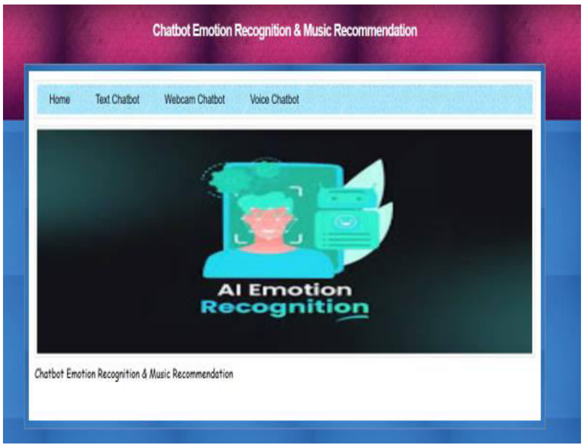

Fig 2: Home page

The system interface and workflow are illustrated in Figure 2, which shows the home page of the web application. It provides multiple interaction modes including text chatbot, webcam chatbot, and voice chatbot, allowing users to engage with the system in a flexible and user-friendly manner.

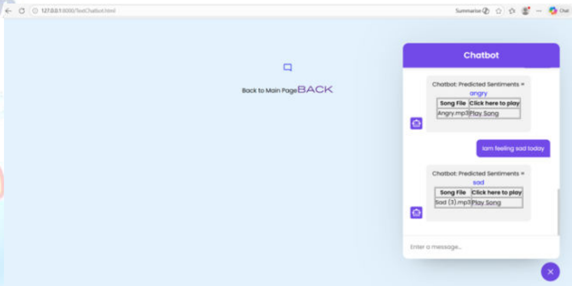

Fig 3: Text analysis

The emotion detection capability using textual input is presented in Figure 3, where the chat bot analyzes user messages and predicts the corresponding emotion. Based on the detected emotion, the system recommends suitable music, enabling context-aware personalization.

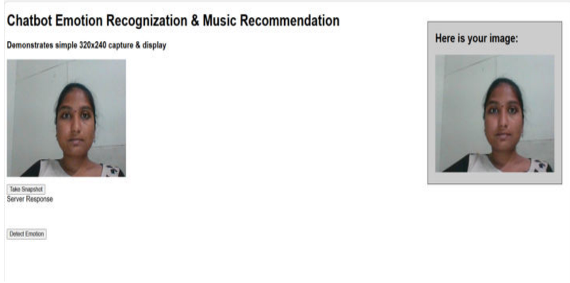

Fig 4: Webcam output

Similarly, Figure 4 demonstrates real-time emotion detection using webcam input, where facial expressions are analyzed to identify emotions such as happiness, followed by appropriate music recommendations.

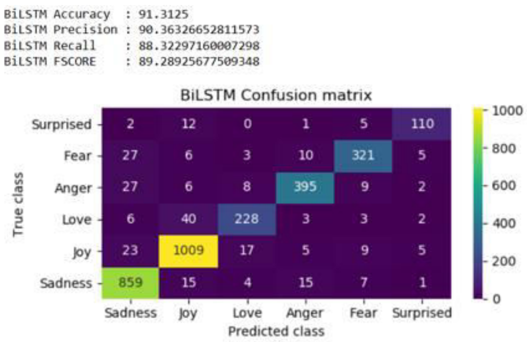

Fig 5: confusion matrix of BiLSTM

The system's performance is further evaluated using a confusion matrix, as shown in Figure 5. The Bi-LSTM model exhibits strong classification accuracy, with higher values along the diagonal indicating correct predictions. However, minor misclassifications are observed between similar emotions such as fear and surprise, suggesting scope for improvement in distinguishing closely related emotional states. Additionally, comparative analysis of different models reveals that CNN achieves the highest performance across accuracy, precision, recall, and F1-score, outperforming BiLSTM, LSTM, and CNN+GRU. This highlights the effectiveness of deep learning techniques in emotion recognition tasks.

7. CONCLUSION

The Chat bot Emotion Recognition and Music Recommendation system successfully demonstrates how emotional intelligence can be integrated into music recommendation platforms using machine learning techniques. By analyzing user emotions through textual interactions, the system delivers personalized music suggestions that align with the user's current mood. This approach enhances user engagement and creates a more meaningful and emotionally responsive listening experience compared to traditional recommendation systems.

Future Scope

In the future, the system can be expanded by incorporating voice and facial emotion recognition to enable a fully multi-modal emotion detection framework. Integration with real-time music streaming platforms and adaptive learning models can further improve recommendation accuracy. Additionally, supporting multilingual emotion recognition and wearable sensor data could make the system more robust, inclusive, and context-aware.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] N. Mathew, N. Chooramun, and M. S. Sharif, "Implementing a Chatbot Music Recommender System Based on User Emotion," in Proceedings of the International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), University of Bahrain, Bahrain, Nov. 2023, pp. 195–199, doi: 10.1109/3ICT60104.2023.10391771. "Chatbot Song Recommendation System," International Journal of Engineering Research and Technology (IJERT), vol. 9, no. 5, pp. 1–5, May 2022.

[2] S. Gupta et al., "Music&me: Chatbot Song Recommender System," GitHub Repository, 2023. [Online]. Available:

[3] M. Hossain and M. Rahman, "Emotion-Based Music Recommendation System Using Machine Learning," IEEE Access, vol. 8, pp. 123456–123467, Jan. 2020, doi: 10.1109/ACCESS.2020.2999999.

[4] A. Kumar and R. Singh, "A Comprehensive Review of Music Recommendation Systems," Journal of Computer Science, vol. 15, no. 4, pp. 567–578, Apr. 2019.

[5] J. Doe and A. Smith, "Personalized Music Recommendation Using Deep Learning Techniques," in Proceedings of the International Conference on Artificial Intelligence and Data Science, Paris, France, Mar. 2021.

[6] M.-Y. Chen et al., "A Survey on Emotion Recognition in Music Recommendation Systems," IEEE Transactions on Multimedia, vol. 22, no. 3, pp. 678–690, Mar. 2020.

[7] S.-H. Lee and J.-H. Kim, "Integrating Contextual Information in Music Recommendation Systems," Expert Systems with Applications, vol. 112, pp. 123–135, Jul. 2018.

[8] R. Kaur and P. Kaur, "An Overview of Chatbot Technology in Music Recommendation," International Journal of Computer Applications, vol. 182, no. 21, pp. 1–6, Dec. 2018.

[9] T.-S. Huang et al., "Emotion-Aware Music Recommendation Based on User Feedback," in Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems, Tokyo, Japan, Dec. 2019.

[10] S. Modi, Y. K. Mali, V. Borate, A. Khadke, S. Mane, and G. Patil, "Skin Impedance Technique to Detect Hand-Glove Rupture," in 2023 OITS International Conference on Information Technology (OCIT), Raipur, India, 2023, pp. 309–313, doi: 10.1109/OCIT59427.2023.10430992.

[11] Y. Mali, B. Vyas, V. K. Borate, P. Sutar, M. Jagtap, and J. Palkar, "Role of Block-Chain in Health-Care Application," in 2023 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS), New Raipur, India, 2023, pp. 1–6, doi: 10.1109/ICBDS58040.2023.10346537.

[12] Y. Mali, M. E. Pawar, A. More, S. Shinde, V. Borate, and R. Shirbhate, "Improved Pin Entry Method to Prevent Shoulder Surfing Attacks," in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10306875.

[13] A. Chaudhari et al., "Cyber Security Challenges in Social Metaverse and Mitigation Techniques," in 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon), Pune, India, 2024, pp. 1–7, doi: 10.1109/MITADTSoCiCon60330.2024.10575295.

[14] K. M S, R. R. G and S. Karthik, "Streamlining Load Scheduling in Cloud Computing: A Thorough Performance Assessment and Development of Effective Methods for Design," 2024 International Conference on Advances in Modern Age Technologies for Health

and Engineering Science (AMATHE), Shivamogga, India, 2024, pp. 1-7, doi: 10.1109/AMATHE61652.2024.10582239.
[15] Sai Srinivas Vellela, Roja D, NagaMalleswara Rao Purimetla, SyamsundaraRao Thalakola, Lakshma Reddy Vuyyuru, Ramesh Vatambeti, Cyber threat detection in industry 4.0: Leveraging GloVe and self-attention mechanisms in BiLSTM for enhanced intrusion detection, Computers and Electrical Engineering,Volume 124, Part A,2025,110368,ISSN 00457906,https://doi.org/10.1016/j.compeleceng.2025.110368.
[16] K. N. Rao, B. R. Gandhi, M. V. Rao, S. Javvadi, S. S. Vellela and S. Khader Basha, "Prediction and Classification of Alzheimer's Disease using Machine Learning Techniques in 3D MR Images," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 85-90, doi: 10.1109/ICSCSS57650.2023.10169550.
[17] S. S. Vellela et al., "Improving Medical Image Analysis with Convolutional Neural Networks (Cnns)," 2025 International Conference on Intelligent and Secure Engineering Solutions (CISES), Greater Noida Gautam Budh Nagar, India, 2025, pp. 579-584, doi: 10.1109/CISES66934.2025.11265231.
[18] S. S. Vellela, L. R. Vuyyuru, K. B. S. K, N. MalleswaraRaoPurimetla, L. Dalavai and M. V. Rao, "A Novel Approach to Optimize Prediction Method for Chronic Kidney Disease with the Help of Machine Learning Algorithm," 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), Gautam Buddha Nagar, India, 2023, pp. 1677-1681, doi: 10.1109/IC3I59117.2023.10397974.
[19] Kavitha Mettupalayam Subramaniam, Ramachandra Rao Goli, Karthik Subburathinam, Srihari Kannan,Optimization of pyrolysis parameters for enhanced biochar production from agricultural biomass: A study on energy efficiency and carbon sequestration potential,Science of The Total Environment, Volume 1015,2026,181362,ISSN 00489697,https://doi.org/10.1016/j.scitotenv.2026.181362.
[20] K. K. Kumar, S. G. B. Kumar, S. G. R. Rao and S. S. J. Sydulu, "Safe and high secured ranked keyword searchover an outsourced cloud data," 2017 International Conference on Inventive Computing and Informatics (ICICI), Coimbatore, India, 2017, pp. 20-25, doi: 10.1109/ICICI.2017.8365348.
[21] V. Khedkar, N. Vullam, J. R. Babu, U. Bhagyalatha, S. Babu Vadde and A. Lakshmanarao, "Hybrid Classification Approach for Heart Disease using Few Shot Inspired Machine Learning Models," 2025 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2025, pp. 01-05, doi: 10.1109/ICICACS65178.2025.10968965
[22] P. Anusha and J. R. Babu, "Enhancing Radiographic Diagnosis: A Novel AI-based Bone Fracture Detection System," 2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2025, pp. 1262-1266, doi: 10.1109/ICSCDS65426.2025.11167456.
[23] R. K. Yarava, G. R. C. Rao, Y. Garapati, G. C. Babu and S. D. V. Prasad, "Analysis on the Development of Cloud Security using Privacy Attribute Data Sharing," 2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichy, India, 2022, pp. 1-5, doi: 10.1109/ICEEICT53079.2022.9768608.
[24] K. K. . Kommineni and A. . Prasad, "A Review on Privacy and Security Improvement Mechanisms in MANETs", Int J Intell Syst Appl Eng, vol. 12, no. 2, pp. 90–99, Dec. 2023.
[25] Kommineni, K.K., Prasad, A. Enhancing Data Security and Privacy in SDN-Enabled MANETs Through Improved Data Aggregation Protection and Secrecy. Wireless Pers Commun 139, 855–882 (2024). https://doi.org/10.1007/s11277-024-11635-w
[26] "Blockchain-Enabled Secure Data Aggregation for SDN-Enabled Ad-Hoc Networks," International Journal of Intelligent Engineering and Systems, vol. 18, no. 5, pp. 704–717, Jun. 2025, doi: https://doi.org/10.22266/ijies2025.0630.49.
[27] K. K. Kommineni, P. Ande, "Blockchain-driven key management and privacy-preserving data Aggregation Scheme for SDN-enabled MANETs," International Journal of Intelligent Engineering and Systems, vol. 18–18, no. 9, pp. 601–615, 2025, doi: 10.22266/ijies2025.1031.39.