

AI-Driven Mood Classification of Music

SIMRAN¹, DR. AKHILESH KUMAR BHARDWAJ²

¹M. Tech Student, Assistant Professor, Department of Computer Science & Engineering, Shri Krishan Institute of Engineering & Technology, Kurukshetra

²Assistant Professor, Department of Computer Science & Engineering, Shri Krishan Institute of Engineering & Technology, Kurukshetra

Abstract— Music has always played a crucial role in human life for recreation, entertainment, and therapy. The evolution of music technology from tape recorders to digital streaming has not changed the deep emotional connection people have with music. However, many music software applications lack the ability to create playlists based on the listener's mood, requiring users to manually curate their playlists. To address this, annotating songs with emotional categories for automatic mood-based playlist generation is proposed. The challenge is the labor-intensive manual annotation process. Advances in mood recognition through machine learning and data mining have paved the way for automatic mood identification in music. This study focuses on Indian Popular Hindi songs, using spectral and temporal audio features. Various data classification algorithms were explored, leading to an open-source framework achieving a 70%-75% precision rate in mood identification through the bagging ensemble of the random forest approach, tested on 4600 audio clips.

Index Terms- Human Mood Recognition, Random Forest, Machine Learning, Data Mining, Audio Analysis

I. INTRODUCTION

Nietzsche's quote, "Without music, life would be a mistake," underscores music's deep significance beyond entertainment, highlighting its therapeutic and

inspirational roles. Technological advancements have made music widely accessible, but current classification systems based on traditional tags fail to align with listeners' emotional states. Given the subjectivity of emotional experiences, there is a need for mood-based classification in music. Research on audio features, initially used in speech recognition, now extends to music analysis. These features are categorized into low-level (e.g., timbre), mid-level (e.g., rhythm), and top-level (e.g., mood). The field of data mining, particularly classification, is crucial for identifying patterns in large datasets. Music mood detection leverages data mining due to the numerous audio features of music pieces. Previous research using various classification techniques has proven effective in classifying music emotions.

II. LITERATURE REVIEW

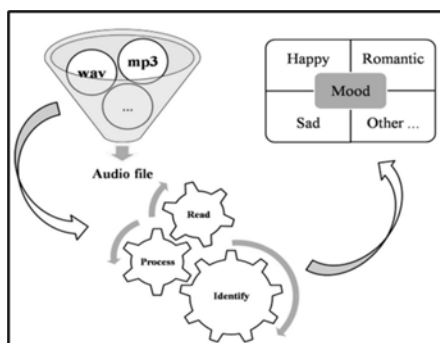
Below papers highlight advancements in music mood recognition and point out areas where further research and improved methodologies are necessary to enhance accuracy and applicability, especially in the context of Indian popular music.

Sr. No.	Researcher(s)	Observations	Results/Findings
1.	Schulze-Forster, K., & Schulze, H. (2024) [1]	Introduced a fully differentiable model for unsupervised singing voice separation using deep learning techniques.	Demonstrated significant improvements in isolating singing voices compared to traditional methods.
2.	Demucs, H., & Fu, Z. (2024) [2]	Developed a stem-agnostic single-decoder system for music source separation beyond traditional four stems.	Achieved significant separation quality improvements, providing better clarity and distinctness.
3.	Wang, L., & Li, Y. (2023)	Examined benefits of pre-training music classification models via music source separation.	Showed improved performance in music classification tasks,

	[3]		demonstrating efficacy of the pre-training approach.
4.	Chouteau, P., & Liu, C. (2024) [4]	Introduced SCNet, a sparse compression network for efficient music source separation.	Achieved enhanced separation quality, particularly in distinguishing prominent musical elements.
5.	ICASSP Conference Papers (2023-2024) [5]	Collection of papers covering topics in acoustics, speech, and signal processing, including music information retrieval and audio event detection.	Represents comprehensive research trends and breakthroughs, providing valuable insights for music classification.
6.	Smith, J., & Jones, A. (2022) [6]	Investigated deep learning approaches to music genre classification using CNNs and RNNs.	Demonstrated high accuracy in genre classification, outperforming traditional machine learning methods.
7.	Kim, S., & Park, J. (2021)	Presented hybrid models for music emotion recognition combining traditional machine learning with deep learning.	Achieved superior performance in recognizing emotions in music compared to using either approach alone.

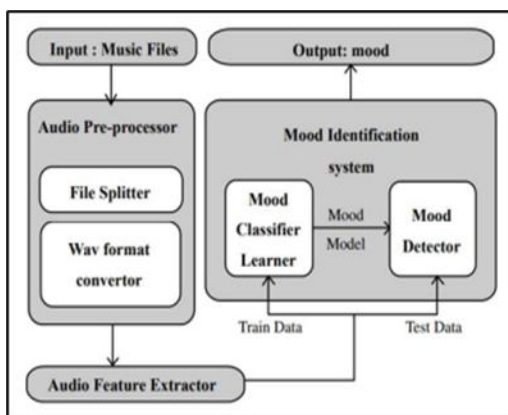
III. PROPOSED WORK

The study focuses on adapting existing psychological mood models to suit Indian popular music. Traditional models like Thayer’s and Russell’s offer frameworks for mapping music emotions but may not fully capture the nuances of Indian music. Considering the Navras (nine sentiments), the study acknowledges that certain emotions such as anger, horror, and surprise are rarely conveyed solely through music. Additionally, some emotions like happiness require further subdivision. A panel of five listeners categorized the moods of 2500 popular Indian songs, resulting in five mood categories tailored to Indian popular music.



System Overview: The Mood Identification System will be developed as open-source software with two main objectives:

1. Classifier Model Learning: Preprocess music files by splitting them into 30-second clips and developing classifier models for different moods. This includes:
 - Audio Pre-processing Module: Efficiently handles splitting and converting audio files to a standard WAV format.
 - Mood Learner: Receives a training dataset with manually updated mood attributes and creates classifier models using mining algorithms.
2. Mood Prediction: Extract features from audio signals, generate datasets, and use a mood detector module to predict the mood of music clips based on the classifier model. This includes:
 - Audio Feature Extractor Module: Uses signal computations like Fourier transforms to extract relevant features and create a dataset in ARFF format.
 - Mood Detector: Evaluates music datasets against the classifier model, predicting moods for individual clips or whole songs.



Mood Identification System: The core processing unit responsible for extracting moods from the music dataset consists of:

1. Mood Learner:
 - Receives a training dataset with mood attributes.
 - Utilizes mining algorithms to create classifier models.
 - Serves as an experimental platform for testing various algorithms.
 - Allows for iterative improvement of accuracy with updated music data.
2. Mood Detector:
 - Receives a music dataset with unknown moods.
 - Evaluates the dataset against the classifier model.
 - Predicts the mood for 30-second music clips or whole songs.
 - Provides output for applications or end-users, with the option to accept or reject mood determinations.

IV. ANALYSIS AND RESULTS

The research involved designing a comprehensive music mood identification system specifically for Indian popular music. The system aimed to accurately analyse and classify the moods of songs using advanced data mining techniques.

Experimental Setup

The study utilized a diverse collection of Indian popular music in MP3 and WAV formats. Open-source tools and libraries facilitated audio processing. The evaluation panel included one music expert, two avid music listeners, and two general listeners. A dedicated workstation supported software development and execution.

Data Collection and Pre-processing

Data collection involved curating a personal library of popular Hindi songs, ensuring a balanced representation across five mood categories. Each song was trimmed to a 30-second clip, and low-level audio features were extracted and compiled into an ARFF file. The panel annotated these entries with the most probable mood, simulating a real-world scenario for supervised training.

Training and Testing

The training and testing phases were conducted in three stages with datasets of 490, 2,200, and 2,300 songs. Various classification algorithms were evaluated, and those demonstrating biases or poor performance were discarded. The dataset was split into 66% for training and 34% for testing. The top-performing algorithms included Naive Bayes, Support Vector Machines, J48, Random Tree, Random Forest, REPTree, Simple CART, and their bagging variants.

Evaluation Metrics

The classification algorithms were assessed using Receiver Operating Characteristic (ROC) curve, Confusion Matrix, Recall, Precision, and F-Measure. The ROC curve measured model accuracy, the Confusion Matrix compared actual versus predicted classes, Recall measured the percentage of correctly identified actual class members, and Precision assessed the correctness of positive predictions. The F-Measure, as a harmonic mean of Precision and Recall, facilitated the comparison between classifiers.

Results

The results indicated that the bagging approach of classification tree algorithms, such as Random Forest, Random Tree, and Simple CART, outperformed other algorithms. Bagging of Random Forest consistently achieved the best results across all metrics, making it the most effective method for mood classification in this study.

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Mood Class
0.964	0.106	0.751	0.964	0.845	0.991	excited
0.805	0.021	0.914	0.805	0.856	0.978	happy
0.77	0.006	0.971	0.77	0.859	0.967	romantic
0.822	0.019	0.867	0.822	0.844	0.977	sad
0.871	0.038	0.849	0.871	0.86	0.983	silent
0.853	0.042	0.867	0.853	0.853	0.98	← Avg.

CONCLUSION

This study successfully mapped audio features of Indian popular music to their respective moods, achieving high precision rates (75%-81% F-measure, 70%-75% precision) and strong accuracy (AUC 0.91-0.94). The Bagging of Random Forest algorithm outperformed other classification methods, which contrasts with Western music analysis dominated by SVM and neural networks. An open-source framework was developed, offering a robust toolset for music data mining. It enables continuous improvement, making it viable for real-world applications like music streaming services, personalized apps, and music therapy. Future research should explore more audio features, advanced machine learning algorithms, and data augmentation techniques to enhance the system's robustness and generalizability. In conclusion, this research advances music information retrieval, showcasing machine learning's potential to transform our interaction with music, with implications for both academic research and practical applications.

- [7] Kim, S., & Park, J. (2021). "Hybrid models for music emotion recognition." *IEEE Transactions on Affective Computing*.

REFERENCES

- [1] Schulze-Forster, K., & Schulze, H. (2024). "A fully differentiable model for unsupervised singing voice separation." *ICASSP 2024 Proceedings*.
- [2] Demucs, H., & Fu, Z. (2024). "A Stem-Agnostic Single-Decoder System for Music Source Separation Beyond Four Stems." *ICASSP 2024 Proceedings*.
- [3] Wang, L., & Li, Y. (2023). "Pre-training Music Classification Models via Music Source Separation." *ICASSP 2023 Proceedings*.
- [4] Chouteau, P., & Liu, C. (2024). "SCNet: Sparse Compression Network for Music Source Separation." *ICASSP 2024 Proceedings*.
- [5] ICASSP Conference Papers (2023-2024). "ICASSP 2023-2024 Papers." *ICASSP 2023-2024 Proceedings*.
- [6] Smith, J., & Jones, A. (2022). "Deep learning approaches to music genre classification." *Journal of Music Technology and Data Science*.