

Classifying Indian Classical Music For Therapeutic Cause Using Lightweight Deep Neural Network

**Arnab Kundu, Bijoy Kar, Parijat Das, Pratingya Sahoo,
Sweta Kumari, Somsubhra Gupta**

Department of Computer Science & Engineering
Swami Vivekananda University

Telinipara, Barasat - Barrackpore Rd Bara Kanthalia
Barrackpore, India

arnabkundu759, bijaykar54321, parijatdas823,
bijaykar54321}@gmail.com, sahoopratingya@gmail.com,
swetakumari272000@gmail.com

Abstract

Indian classical music, known for its emotional depth and complex melodies, is increasingly recognized for its therapeutic benefits. Genres like Bhajan, Thumri, Ghazal, and Sufi, paired with specific ragas such as Yaman (energizing) and Bhairavi (calming), offer unique effects like stress relief, emotional regulation, and introspection. A proposed lightweight deep learning model classifies genres and ragas by analysing key audio features such as rhythm, melody, pitch, and tonal patterns. The model captures the nuances of Indian music, ensuring precise therapeutic recommendations by accurately identifying genre-raga combinations for optimal psychological and emotional benefits.

Keywords : Indian Classical Music, Deep Neural Network, Classification, Therapeutic Cause

1. Introduction

Indian classical music, known for its emotional depth and therapeutic value, can realign vibrations in the body, aiding in healing and restoring balance. Music stimulates the pituitary gland, releasing neurochemicals that boost brain activity, emotional well-being, and healing. Techniques like **Raga Chikitsa** use specific ragas to evoke moods and support

therapy. This research combines traditional music therapy with modern technology through a lightweight deep neural network (DNN) model. The DNN effectively classifies subtle genre and raga variations, ensuring accurate therapeutic applications. By analyzing key features like rhythm, pitch, and frequency, the model addresses issues ranging from stress to chronic conditions, bridging the gap between Indian classical music's traditional healing methods and advanced machine learning. The study aims to create an efficient, adaptable tool for music therapy, contributing to global mental and emotional health advancement

2. Previous Work

Music, a universal language, evokes emotions and influences behavior, transcending barriers. Its therapeutic potential impacts both living and non-living entities, with Indian classical music offering profound healing benefits. Understanding its mechanisms helps harness music's power for personal and global well-being

Dr. Sangeeta Venkatesh's research on **Raga Therapy** highlights how ragas like Yaman reduce stress and promote calmness, aiding in conditions like insomnia and anxiety. Similarly, Dr. Subhadra Chaudhary found in her paper **The Impact of Raga Bhairav on Anxiety Reduction** that this raga, known for its meditative quality, effectively lowers anxiety levels and can help manage high blood pressure.

Dr. T.V. Sairam emphasizes in Raga Therapy how specific ragas influence the body's energy centers, aiding in healing. **Raga Darbari**, for instance, helps relieve pain and induces relaxation. Similarly, Dr. Shantala Hegde's research reveals that Indian classical music supports individuals with depression by fostering mental engagement and emotional release. These studies showcase its growing role in holistic therapy for emotional and physical well-being.

Historical texts like Sangita Sudha and the Saraswati Mahal records (1684–1711) highlight the enduring belief in Indian classical music's therapeutic power, documenting its use for emotional and psychological healing.

A 2020 paper published in the **Indian Journal of Applied Research (IJAR)**, titled “**Therapeutic Effects of Ancient Indian Classical Music.**” explores its emotional and psychological benefits, reaffirming its historical role in therapy.

The ancient manuscript Raga Chikitsa discusses using ragas like Darbari Kanada and Khamaj to treat mental tension and hysteria. Modern studies, such as one in the Manipal Journal of Medical Sciences, continue to explore the therapeutic effects of ragas, offering a contemporary perspective on these traditional healing practices.

Thaut, M. H., & Hoemberg, V. (2014) provided foundational insights into Neurologic Music Therapy in their Handbook of Neurologic Music Therapy. Their work highlights the therapeutic potential of music in treating neurological disorders, emphasizing the integration of music and neuroscience for enhanced clinical outcomes. This resource has been instrumental in shaping current therapeutic practices and research in the field.

Recent research on music genre classification using deep learning has explored various approaches:

- 1. CNN-Based Models:** A deep CNN model with dense layers, dropout, and ReLU activation functions, optimized with Adam and sparse categorical cross-entropy, showed promising results in genre classification.
- 2. Hybrid CNN-LSTM Approach:** Combining CNNs for feature extraction and LSTMs for sequential learning improves classification by capturing both spatial and temporal features of music.
- 3. Multilingual DNNs:** A multilingual DNN framework, adapted from automated speech recognition, achieved a 93.4% accuracy on the GTZAN dataset by using pre-training data from related genres and tackling underrepresented genres.
- 4. DNNs with MFCC Features:** A DNN model using Mel-Frequency Cepstral Coefficients (MFCC) features achieved 97.8% accuracy in classifying classical, pop, rock, and electronic genres, outperforming traditional machine learning methods.

5. Architectural Variations: Models with different architectures, ranging from shallow to deep networks, tested with different optimizers (Adam, SGD) and epochs, demonstrated that layer complexity and optimizer choice significantly impact performance.

6. Spiking Neural Networks (SNNs): Various methods for optimizing an SNN model showed that varying the number of dense layers, epoch count, and optimizer choice (Adam, SGD, RMSprop) impacted the classification accuracy.

7. Raga Identification: Feature extraction techniques like MFCC, LPC, and chroma-based features are critical for raga identification, with clustering and machine learning models such as K-means, KNN, SVM, and HMM being commonly used.

Additionally, studies highlight the therapeutic effects of music, particularly Indian classical music (ICM) and ragas, which are used for health benefits, including relaxation and emotional well-being.

3. Research Methodologies

This study focuses on classifying Indian Classical Music for therapeutic purposes, aiming to address mental health issues and benefit the general public. The primary research question is how to classify Indian Classical music for therapeutic use using a lightweight deep learning neural network. The study utilizes appropriate data collection methods and public data sources, employing standard data analysis techniques to ensure reliable and generalizable results. Key features are identified for classification, and the model building process incorporates best practices and state-of-the-art technologies. By combining quantitative and qualitative methods, the study aims to provide a comprehensive understanding of classifying Indian Classical music for therapeutic purposes.

3.1. Data Collection & Processing

A large number music data is collected from various sources with the “*fair use law in copyright act*” in mind. In total more than 500 songs are collected from these four genres –

Bhajan, Garhwali, Gazal and Sufi. To capture all the short time duration unique features in the music which belong to different genres all the music is segmented into five seconds clips. This technique also expands the dataset to provide more data for a better training. To extract data from the audio files python library *librosa* is used. First, all the audio data is segmented into five seconds clips, then all the audio features are extracted from each clip. The extracted features are *audio_name*, *peak_frequency*, *peak_pitch*, *power*, *avg_spectral_centroid*, *max_spectral_centroid*, *avg_spectral_rolloff*, *max_spectral_rolloff*, *avg_zero_crossing_rate*, *max_zero_crossing_rate*, *avg_rms*, *max_rms*, *avg_chroma*, *max_chroma*, *avg_mfccs*, *max_mfccs*, *tempo*, *songtype*. Among these audio features spectral centroid, spectral rolloff, zero crossing rate, rms, chroma and mfccs are non-scaler data (1D or 2D data). These features are dimensionality reduction using basic statistical rules like mean and max. This technique helps to achieve a simpler dataset and an appropriate light weight model architecture. In the data extracting process both the mean and the max are extracted of the non-scaler features, because in the data analysis phase only the appropriate column will be taken from each feature.

3.2. Data Analysis & Feature Selection

At first the target column (*songtype*) is *OneHotEncoded* into four different columns representing four genres of music. Then the unnecessary features like *audio_name* and *songtype* are removed from the dataset. A correlation matrix and a heatmap is created from the updated dataset. The dataset contains few features which are highly correlated with each other, like *avg_spectral_rolloff* and *max_spectral_rolloff*, *avg_zero_crossing_rate* and *max_zero_crossing_rate*. In these cases, these highly correlated features or variable cause collinearity, which may affect the model performance, so one of them is removed from the dataset based on the correlation with the targets.

3.3. Model Architecture

A neural network architecture is the design and structure of a neural network, defining the arrangement of layers, neurons,

and connections between them. A neural network can have one input layer, one output layer and multiple hidden layers. A neural network is very versatile and can recognize patterns in data. This model contains 8 layers in total.

- **Input Layer:** The input layer is configured to accept features with a shape of. This ensures compatibility with the training dataset.

- **Normalization Layer:** A normalization layer standardizes the input features by scaling them to a standard range, improving the stability and efficiency of the training process.

- **First Hidden Layer:** This layer contains 64 neurons, each using the Scaled Exponential Linear Unit (SELU) activation function. The kernel weights are initialized using a random normal distribution.

- **Second Hidden Layer:** A dense layer with 128 neurons is introduced, also utilizing the SELU activation function and random normal weight initialization.

- **Third Hidden Layer:** This layer comprises 256 neurons, again employing the SELU activation function with random normal initialization.

- **Fourth Hidden Layer:** A dense layer with 128 neurons, activated by SELU, follows the structure of the previous layers.

- **Fifth Hidden Layer:** This layer contains 64 neurons with SELU activation and random normal initialization.

- **Output Layer:** The final layer is a dense layer with a number of neurons equal to the number of target classes in the dataset. It uses the *softmax* activation function to produce probability distributions over the target classes.

Loss Function: The loss function is a crucial component of a neural network, serving as a measure of how well the model's predictions align with the true target values. It guides the optimization process by quantifying the error, which the model minimizes during training. The *Categorical Crossentropy* loss function was selected for this model as it is well-suited

for multi-class classification tasks where the target variable is one-hot encoded. This loss function calculates the difference between the true class labels and the predicted probability distribution generated by the model's *softmax* output. By penalizing incorrect predictions proportionally to their confidence, *Categorical Crossentropy* encourages the model to produce high probabilities for the correct classes while minimizing the likelihood of incorrect classes. This makes it an effective choice for ensuring robust learning and achieving accurate classification in multi-class settings.

$$Loss = - \sum_{i=1}^N y_i \log(p_i)$$

Where N is the number of classes. y_i is the true label for class i (in one-hot encoded form, it is either 0 or 1). p_i is the predicted probability for class i (the output from the model's *softmax* function).

Activation Function: For this model SELU is used over RELU as an activation function. SELUs, or Scaled Exponential Linear Units, are activation functions that induce self-normalization. SELU network neuronal activations automatically converge to a zero mean and unit variance.

$$SELU(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ a(e^x - 1) & \text{if } x < 0 \end{cases}$$

Where λ is a predefined scaling parameter, and a is another predefined parameter typically derived from fixed-point equations that ensure self-normalization. Specifically, a and λ are defined to ensure that the mean and variance remain invariant across layers, which are crucial for deep neural networks to avoid the vanishing or exploding gradient problems.

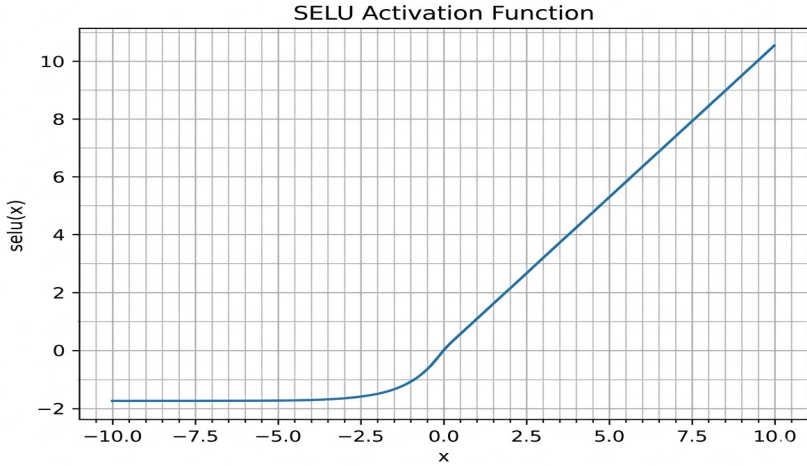


Figure 1: SELU activation graph

The Scaled Exponential Linear Unit (SeLU) activation function was chosen for this model due to its self-normalizing properties, which help maintain stable mean and variance of activations across layers during training. This eliminates the need for additional normalization techniques like batch normalization, simplifying the architecture while ensuring efficient gradient flow. SeLU's ability to mitigate vanishing and exploding gradients makes it particularly well-suited for deeper networks, as it facilitates stable and consistent training dynamics. Additionally, its smooth, differentiable nature and suitability for deep learning tasks make it an ideal choice for achieving robust performance in this model.

Model Optimizer: The Nadam optimizer (Nesterov-accelerated Adaptive Moment Estimation) is an extension of the Adam optimizer that integrates Nesterov momentum. It combines the benefits of adaptive learning rates and momentum-based gradient updates, making it particularly effective for deep learning tasks.

$$\theta_{t+1} = \theta_t - \frac{\eta}{(\sqrt{\hat{v}_t} + \epsilon)} (\beta_1 m_t + \frac{(1 - \beta_t) g_t}{1 - \beta_1^t})$$

Nadam (Nesterov-accelerated Adaptive Moment Estimation) optimizer was selected for this model due to its ability to combine the benefits of adaptive learning rates and momentum-based gradient updates. Nadam extends the Adam optimizer by incorporating Nesterov momentum, which provides a lookahead mechanism, enabling faster convergence and more efficient navigation of the optimization landscape. This makes it particularly effective for complex models with multiple layers, where precise and adaptive gradient updates are crucial. Its smooth convergence properties and ability to handle sparse gradients make Nadam a robust choice for achieving optimal performance in this deep learning model.

Model Weight Initialization: The random normal initializer was used in this model to initialize the weights with values drawn from a normal distribution, which aligns well with the requirements of the *SeLU* activation function. Proper weight initialization is crucial for maintaining stable gradients and activations during training, and the random normal initializer ensures that the weights have an appropriate scale to support *SeLU*'s self-normalizing properties. This choice helps the model converge efficiently by avoiding issues like vanishing or exploding gradients, making it particularly suitable for deep architectures.

$$W \sim \mathcal{N}(\mu, \sigma^2)$$

Where W is the initialized weight. μ is the mean of the distribution. σ is the standard deviation, which is the square root of the variance.

Learning Rate Scheduler: A custom learning rate scheduler was implemented using the 1-cycle learning rate policy to enhance the training process. The scheduler maintains a constant learning rate of 0.0010 for the initial 100 epochs, allowing the model to converge efficiently during the early stages of training. After 100 epochs, the learning rate decays exponentially by a factor of $e^{0.1}$ per epoch, enabling fine-tuning and stabilizing the optimization process in later stages. This approach balances rapid initial learning with a gradual reduction in learning rate, which helps avoid overshooting the minimum and facilitates convergence to an optimal solution.

$$LR = \begin{cases} 1 & \text{if } \epsilon \leq 100 \\ 1e^{-0.10} & \text{if } \epsilon > 100 \end{cases}$$

Where LR is the calculated learning rate for each epoch (ϵ).

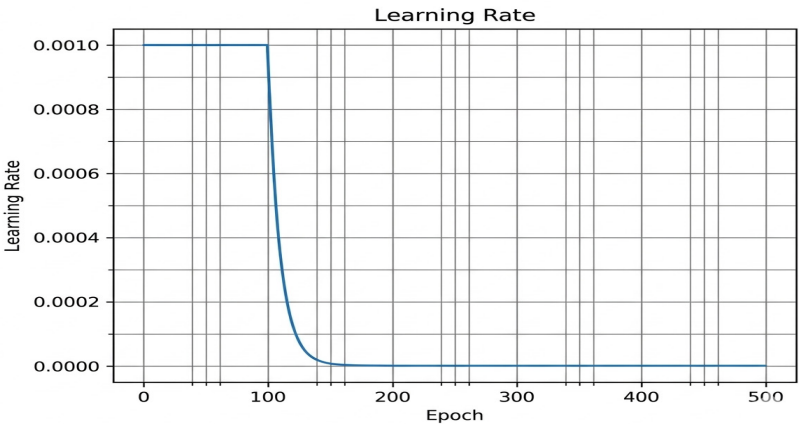


Figure 2: Learning Rate Graph

3.4. Model Performance

The performance of the model was evaluated using precision, recall, and F1-score across four classes: **Garhwali**, **Ghazal**, **Bhajan**, and **Sufi**.

The results for each class are summarized below:

- **Garhwali:** The model achieved a precision of **0.80**, recall of **0.80**, and F1-score of **0.80**, indicating a balanced and accurate performance for this class with high prediction quality and coverage.
- **Ghazal:** Precision was **0.70**, recall **0.68**, and F1-score **0.69**. While the model performed reasonably well on this class, there is room for improvement in both precision and recall.
- **Bhajan:** This class saw a precision of **0.75**, recall of **0.75**, and F1-score of **0.75**, demonstrating consistent and reliable predictions.

- **Sufi:** The model’s performance on Sufi was slightly lower, with precision **0.68**, recall **0.69**, and F1-score **0.69**, suggesting that the model may have slightly struggled to differentiate this class as effectively as the others.

Overall, the model achieved an **accuracy of 0.73**, with a **macro average** and **weighted average** of **0.73** for precision, recall, and F1-score. The macro average reflects the model’s general performance across all classes, treating each class equally, while the weighted average considers the support (number of samples) for each class, providing a more balanced overall metric.

	Precision	Recall	F1-Score	Support
Garhwali	0.80	0.80	0.80	3149
Ghazal	0.70	0.68	0.69	2305
Bhajan	0.75	0.75	0.75	2749
Sufi	0.68	0.69	0.69	3047
Accuracy			0.73	11250
Macro Avg	0.73	0.73	0.73	11250
Weighted Avg	0.73	0.73	0.73	11250

Table 1: Model Classification Report

Confusion Matrix Analysis: The diagonal elements (2519, 1569, 2068, 2109) represent correct predictions, while off-diagonal elements show misclassifications (from *Figure 3*).

Notable Observations:

- For Garhwali it shows excellent performance with 2519 correct predictions and lowest misclassification rates across other categories (142, 197, and 291).
- For Ghazal it shows the confusion with other categories, particularly with 370 instances being misclassified as Sufi.
- For Bhajan it shows some confusion with other categories, particularly 309 instances misclassified as Sufi.
- For Sufi relatively consistent misclassification pattern across other categories.

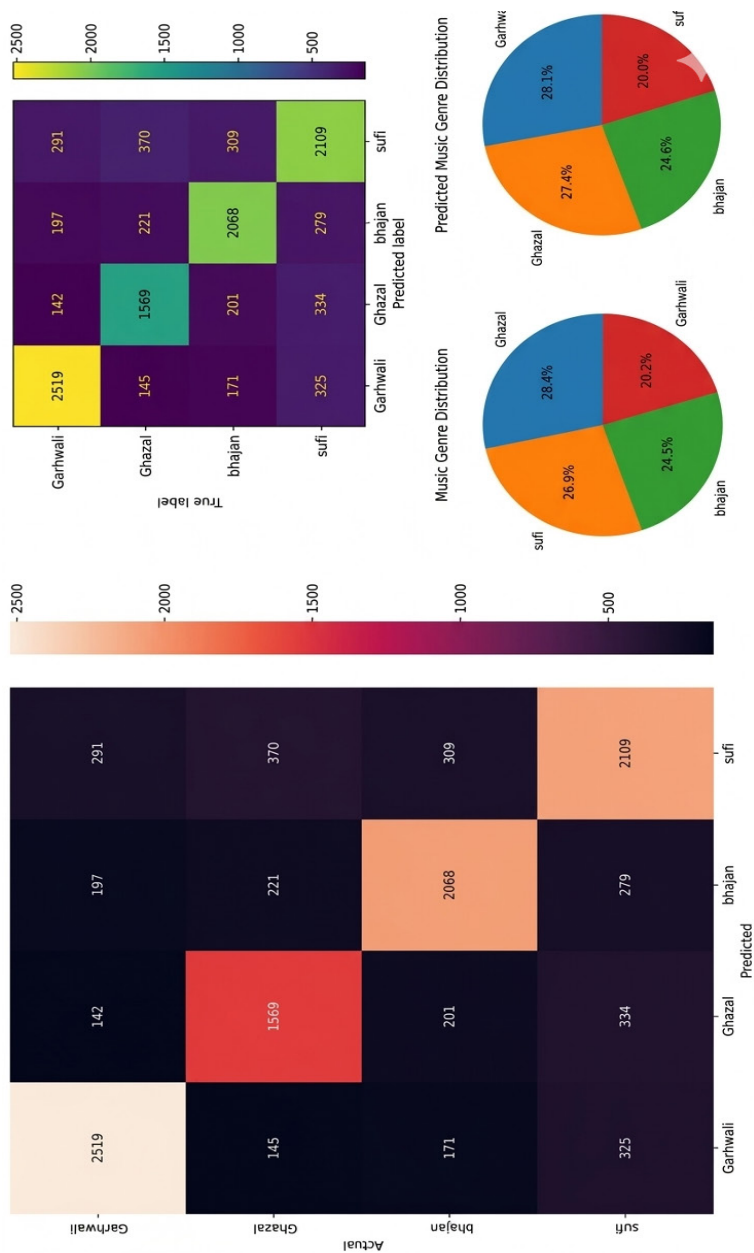


Figure 3: Heatmap of Confusion Matrix

Areas for Improvement:

- The model shows some confusion between Ghazal and Sufi categories.
- There's a consistent pattern of misclassification into the Sufi category across all other classes.

Analysis Conclusion: While the model demonstrates robust overall performance, particularly for Garhwali classification, there remains room for improvement in distinguishing between Ghazal and Sufi forms. Future work should focus on enhancing feature extraction methods to better capture the unique characteristics of each musical form, particularly for the more frequently confused categories.

Conclusion

This study presents a lightweight deep neural network model for classifying Indian classical music genres with therapeutic applications. The model achieved notable performance metrics, with an overall accuracy of 73% across four distinct musical categories (Garhwali, Ghazal, Bhajan, and Sufi). The Garhwali classification showed particularly strong results with 80% accuracy, while other genres demonstrated promising but varying levels of performance (Ghazal: 69%, Bhajan: 75%, Sufi: 69%). The model's architecture, incorporating SELU activation functions and Nadam optimization, proved effective for capturing the nuanced characteristics of Indian classical music. The implementation of a custom learning rate scheduler and strategic feature selection contributed to the model's efficiency while maintaining computational lightweight properties.

However, several challenges remain, particularly in distinguishing between Ghazal and Sufi genres. Future research directions should focus on:

1. Enhancing feature extraction methods to better capture genre-specific characteristics
2. Expanding the dataset to include more diverse examples within each genre
3. Investigating advanced architectures that maintain

lightweight properties while improving cross-genre discrimination

4. Incorporating raga-specific features to better align with therapeutic applications

The successful development of this classification system represents a significant step toward integrating traditional Indian classical music therapy with modern machine learning approaches, potentially benefiting mental health practitioners and patients alike.

Acknowledgments

We would like to express our heartfelt gratitude to our project supervisor, Prof. Somsubhra Gupta, for his constant guidance, valuable suggestions, and encouragement, which were instrumental in the successful completion of our minor project, "Classifying Indian Classical Music for Therapeutic Cause Using Lightweight Deep Neural Network." His enthusiasm for research and teaching inspired us to approach challenges with a positive and adventurous mindset.

References

Tarannum Shaikh, "Music Genre Classification Using Neural Network," Academia.edu. [Online]. Available: https://www.academia.edu/109440768/Music_Genre_Classification_Using_Neural_Network. [Accessed: 15-jan-2025].

Kai Middlebrook, "Deep Neural Network Architectures For Music Genre Classification," Academia.edu. [Online]. Available: https://www.academia.edu/116110340/Deep_Neural_Network_Architectures_For_Music_Genre_Classification. [Accessed: Jan. 15, 2025].

Jia Dai, "Multilingual Deep Neural Network for Music Genre Classification," Academia.edu. [Online]. Available: https://www.academia.edu/98378831/_Multilingual_Deep_Neural_Network_for_Music_Genre_Classification. [Accessed: Jan. 15, 2025].

G. Jawaharlalnehru, "Music Genre Classification using Deep Neural Networks," Academia.edu. [Online]. Available:

https://www.academia.edu/37077574/Music_Genre_Classification_using_Deep_Neural_Networks. [Accessed: Jan. 15, 2025].

Anahitaa, "Music Genre Classification using Optimized Sequential Neural Network," Academia.edu. [Online]. Available: https://www.academia.edu/49204623/Music_Genre_Classification_using_Optimized_Sequential_Neural_Network. [Accessed: Jan. 15, 2025].

Dr. Shambhavi Das, "Therapeutic Aspects of Indian Classical Music," Academia.edu. [Online]. Available: https://www.academia.edu/27523247/Therapeutic_Aspects_of_Indian_Classical_Music. [Accessed: Jan. 15, 2025].

S. Subrahmanya, "Raga Chikitsa: An ancient healing art for treating mental disorders," *Manipal Journal of Medical Sciences*, vol. 6, no. 2, pp. 144-149, 2017. [Online]. Available: <https://impressions.manipal.edu/cgi/viewcontent.cgi?article=1144&context=mjms#:~:text='Raga%20Chikitsa'%2C%20an%20ancient,treating%20mental%20disorders%20%5B9%5D>. [Accessed: 15-Jan-2025].

V. Gandhe and M. Tare, "Therapeutic Effects of Ancient Indian Classical Music," *Indian Journal of Applied Research*, vol. 10, no. 8, pp. 41-42, Aug. 2020. [Online]. Available: [https://www.worldwidejournals.com/indian-journal-of-applied-research-\(IJAR\)/recent_issues_pdf/2020/August/therapeutic-effects-of-ancient-indian-classical-music_August_2020_1596272132_1811810.pdf](https://www.worldwidejournals.com/indian-journal-of-applied-research-(IJAR)/recent_issues_pdf/2020/August/therapeutic-effects-of-ancient-indian-classical-music_August_2020_1596272132_1811810.pdf). [Accessed: 15-Jan-2025].

S. Sundar and P. N. Parmar, "Music Therapy: Bridging Traditional Healing System and Modern Science," *Annals of SBV*, vol. 5, no. 2, pp. 33-34, Jul.-Dec. 2016. [Online]. Available: <https://aos.sbvjournals.com/doi/AOS/pdf/10.5005/jp-journals-10085-5210>. [Accessed: 15-Jan-2025].

S. Kumar, "The Role of Bhajans in Spiritual Healing," *Reality Pathing*, Sep. 28, 2024. [Online]. Available: <https://realitypathing.com/the-role-of-bhajans-in-spiritual-healing/>. [Accessed: 15-Jan-2025].

R. Raghunathan, "Thumri, Ghazal, and Modernity in Hindustani Music Culture," *CUNY Academic Works*, 2021. [Online]. Available: https://academicworks.cuny.edu/cgi/viewcontent.cgi?params=/context/jj_pubs/article/1320/&path_info=Thumri_Ghazal_and_Modernity_in_Hindustani_Music_Culture.pdf. [Accessed: 15-Jan-2025].

S. N. Patel, "The Benefits of Listening to Religious Music and Improving Mental Health," *The Ismaili*, Jan. 15, 2025. [Online]. Available: <https://the.ismaili/uk/the-benefits-listening-religious-music-and-improving-mental-health>. [Accessed: 15-Jan-2025].

S. Sharma and A. Verma, "Effect of Music on Animal Behavior: A Review," *ResearchGate*, 2019. [Online]. Available: https://www.researchgate.net/publication/335224477_Effect_of_Music_on_Animal_Behavior_A_Review. [Accessed: 15-Jan-2025].

Mark D. Wilkinson, "Title of the Article," *Nature*, [Online]. Available: <https://www.nature.com/articles/sdata201618>

A. S. Reddy and B. Kumar, "Therapeutic Aspects of Indian Classical Music," *ResearchGate*, 2019. [Online]. Available: https://www.researchgate.net/publication/334559610_Therapeutic_Aspects_of_Indian_Classical_Music. [Accessed: 15-Jan-2025].

A. K. Sharma et al., "The Role of Music Therapy in Mental Health," *PMC*, 2020. [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8566759/>. [Accessed: 15-Jan-2025].

"The Benefits of Music for Mental Health," *The Jed Foundation*, [Online]. Available: <https://jedfoundation.org/resource/how-music-can-improve-your-mental-health/>. [Accessed: 15-Jan-2025].

8 S. K. Yadav and M. Kumar, "Effect of Music on Plants - An Overview," *ResearchGate*, 2016. [Online]. Available: https://www.researchgate.net/publication/291086163_Effect_of_Music_on_Plants_-_An_Overview. [Accessed: 15-Jan-2025].