Indian Music Knowledge Evaluation using Generative AI

K.V. Srilakshmi Asharani[1], Dr. M.V.R. Jyothisree[2], Harini Anand Cherkuri[3], Koottala Navya[4], Bathula Rishishwar Goud[5]

[1]Assistant Professor, Dept. of CSE, MVSR Engineering College, Hyderabad, India [2]Associate Professor, Dept. of CSE, MVSR Engineering College, Hyderabad, India
[3][4][5] IV Year B.Tech CSE, MVSR Engineering College, Hyderabad, India

Abstract

This paper presents an innovative AI-powered platform designed to evaluate knowledge in Indian classical music, seamlessly integrating modern technology with traditional musical pedagogy. The system supports both Carnatic and Hindustani styles and utilizes OpenAI's GPT for intelligent question generation and contextual evaluation, along with Whisper for real-time audio transcription. It accommodates a range of input formats, including text, image, audio, and live singing, thereby enabling comprehensive and multimodal assessment. Unlike conventional systems that rely on static datasets, this platform dynamically interprets and evaluates user responses using generative AI models. Developed using Java-based web technologies, the platform offers a scalable, interactive, and culturally grounded solution for enhancing music education in the digital era.

Keywords

Indian Classical Music, Carnatic, Hindustani, GPT, Whisper, Generative AI, Music Education, Java Web Application

1. Introduction

Indian classical music, with its origins dating back thousands of years, stands as a profound testament to the cultural and spiritual heritage of the Indian subcontinent. Broadly categorized into two distinct traditions—Carnatic, primarily practiced in South India, and Hindustani, dominant in North India—this musical form is characterized by intricate melodic (raga) and rhythmic (tala) frameworks. Mastery of these systems typically requires years of rigorous, personalized instruction under the mentorship of a guru.

Despite its richness, learning Indian classical music presents significant challenges in the modern era. The availability of expert mentors, structured evaluation mechanisms, and consistent feedback has become increasingly limited—especially for learners residing in geographically dispersed or resource-constrained environments. Furthermore, the subjective nature of musical interpretation often complicates standardized assessment and personalized-learning-paths.

Recent advancements in artificial intelligence (AI) and machine learning have begun to revolutionize the landscape of education, including music pedagogy. Generative AI, in particular, offers immense potential for automating, personalizing, and scaling educational experiences across diverse domains. Within this context, the integration of AI tools like OpenAI's GPT and Whisper enables intelligent evaluation of musical knowledge, audio interpretation, and real-time feedback mechanisms. This research presents a novel web-based platform that leverages generative AI to evaluate users' knowledge of Indian classical music. The system supports multiple input modes—text, image, audio, and live singing—and provides contextual, culturally grounded feedback.

2. Related Work

Literature Survey

The fusion of artificial intelligence and Indian classical music has seen significant progress in recent years. Researchers have explored a variety of deep learning architectures, pattern recognition techniques, and audio processing tools to automate and enhance music education and analysis.

Madhusudhan and Chowdhary (2024) introduced DeepSRGM, an LSTM-RNN based system for classifying and ranking musical sequences in Carnatic music, demonstrating effective automated raga identification. Complementing this, Singh and Arora (2024) proposed an explainable deep learning approach combining CNN-LSTM architectures with interpretability tools such as GradCAM++ and SoundLIME to enhance transparency in raga classification.

Ravikumar (2024) developed DAIRHuM, a platform aligning AI-generated outputs with human musical judgments, thereby improving reliability and pedagogical trust in AI systems for Carnatic music. Hebbar and Jagtap (2022) conducted a comparative analysis of audio preprocessing techniques like MFCCs and chromagrams, highlighting their impact on the accuracy of raga recognition algorithms.

In the educational domain, Jamshidi et al. (2021) presented a review of machine learning-enabled music teaching platforms, emphasizing the democratization of learning through intelligent feedback mechanisms. Velankar et al. (2021) focused on melodic pattern recognition by using motif detection to simulate the human-like identification of ragas.

Ventura (2019) explored the application of AI in inclusive education, particularly for dyslexic learners, showcasing how adaptive AI tools in music education can foster skill development in diverse learner populations. Gulati et al. (2016) proposed the Time-delayed Melody Surfaces approach, which maps note transitions over time for improved raga classification.

Further, Gulati, Serra, and colleagues (2014) contributed to large-scale melodic pattern mining in Indian art music archives, aiding AI-based music analysis. Finally, Koduri et al. (2012) addressed **automatic tonic (Sa) identification**, a foundational element for pitch-based systems in Indian classical music evaluation.

S.No	Title	Authors	Year
1	DeepSRGM: Sequence Classification and Ranking in Indian Classical Music	Madhusudhan & Chowdhary	2024
2	Explainable DL Analysis for Raga Identification	Singh & Arora	2024

S.No	Title	Authors	Year
3	DAIRHuM: Aligning AI with Human Judgments	Ravikumar	2024
4	Comparison of Audio Preprocessing Techniques	Hebbar & Jagtap	2022
5	ML-based Music Teaching Platform Review	Jamshidi et al.	2021
6	Melodic Pattern Recognition for Raga ID	Velankar et al.	2021
7	AI in Music Education for Dyslexia	Ventura	2019
8	Time-delayed Melody Surfaces	Gulati et al.	2016
9	Mining Melodic Patterns in Indian Music	Gulati, Serra et al.	2014
10	Automatic Tonic Identification	Koduri et al.	2012

Table 1. Summary of Related Works

3. Proposed Work

The proposed system is a web-based platform designed to evaluate a user's knowledge of Indian classical music through the integration of artificial intelligence. It caters to both *Carnatic* and *Hindustani* traditions, thereby offering a holistic approach to Indian music education. The platform has been developed using a full-stack architecture comprising Java, JSP, HTML, CSS, and MySQL, and is tightly integrated with OpenAI's GPT and Whisper APIs to facilitate advanced AI-driven evaluation.

Unlike conventional systems that rely on static question banks or predefined datasets, this platform dynamically generates and assesses user responses using the GPT language model. It supports multiple input formats across four evaluation levels, including text-based responses, image recognition tasks, audio identification, and live singing performance.

In the final level, users are required to sing a raga phrase, which the system captures and processes using Whisper for audio transcription. The transcribed lyrics are then analyzed in conjunction with the pitch data extracted through a Java-based algorithm. This dual evaluation—comparing both the lyrical accuracy and swara alignment—is performed by GPT against canonical raga structures, enabling personalized, real-time feedback.

This novel approach enables a context-aware and data-agnostic evaluation mechanism, thereby offering learners meaningful insights into their performance without requiring pre-labelled datasets. The system exemplifies how generative AI can be effectively applied to preserve and modernize the pedagogy of Indian classical music.
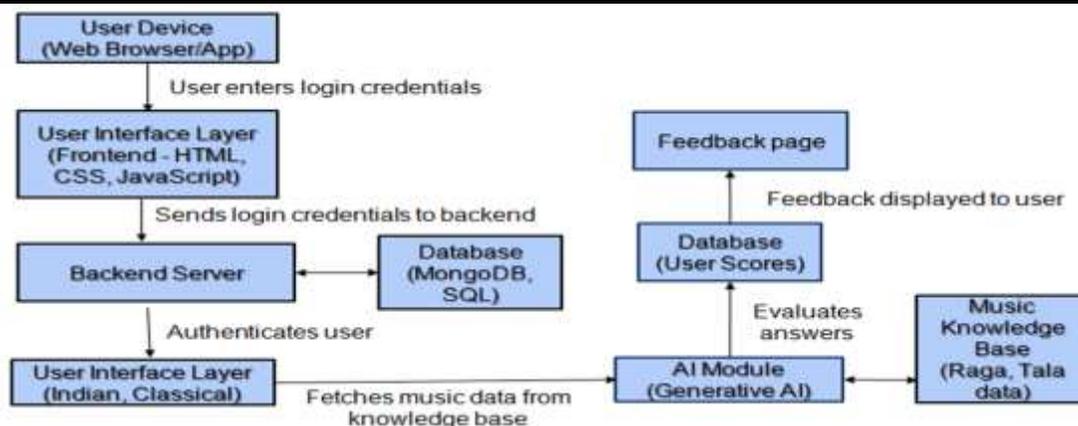
Fig1. Proposed Architecture

The system follows a modular architecture comprising a user interface, backend logic, AI integration, and a centralized database. The frontend is developed using HTML, CSS, and JSP to support smooth navigation and user interaction. The backend, implemented in Java, manages user authentication, level routing, and request handling for AI-based evaluation.

A key component is the integration with OpenAI APIs for processing responses and generating feedback. The system uses a MySQL database to store user details, responses, and scores. Results are rendered dynamically on the interface, and eligible users are provided the option to generate a certificate. This architecture ensures seamless connectivity between user actions and AI-driven evaluation while maintaining efficiency and modularity.

4. Implementation

The Indian Music Knowledge Evaluation system using Generative AI is implemented as a full-stack web application designed to evaluate users' knowledge of Indian classical music. The application leverages AI models from OpenAI, including GPT-3.5 and Whisper, to assess various levels of musical knowledge, ranging from multiple-choice questions (MCQs) to pitch recognition and audio transcription. The system is developed using Java technologies for the backend, MySQL for database management, and HTML, CSS, and JSP for the frontend.

4.1 AI Model Integration

The core of the evaluation process relies on GPT-3.5 and Whisper to interact with users and evaluate their responses. The models are integrated into the backend as follows:

Level 1: MCQs Generation and Evaluation

➢ GPT-3.5 is used to generate multiple-choice questions (MCQs) based on Indian classical music topics, including ragas, talas, and music theory.
➢ Upon receiving user responses, GPT-3.5 evaluates the answers and provides real-time feedback.

Level 2 & 3: Image and Audio Response Interpretation

➢ In these levels, users may submit image or audio responses, such as identifying ragas from visual images or singing a raga.
➢ Whisper transcribes the audio input to text, and GPT-3.5 interprets the text to validate user responses.
➢ Images are analyzed for relevant visual cues related to ragas and musical notation.

Level 4: Pitch Evaluation

➢   The backend uses a Java-based pitch extraction algorithm to analyze the user's audio input.

➢   The pitch is then compared against a reference raga's swara pattern using GPT-3.5 to evaluate the note-by-note accuracy.

➢   The system measures the pitch deviations and compares them with the reference raga's intended swaras.

Textual Comparison (Levenshtein Distance):

For quiz levels involving textual responses (e.g., identifying lyrics or music-related facts), the Levenshtein Distance algorithm is employed to compare user inputs with correct answers, allowing for minor typos and variations in spelling.

4.2 Audio and Image Management

The system stores and organizes user-submitted audio and image files into structured server-side folders. Each file is associated with a specific quiz level and corresponds to the correct answers for validation. Folder-based naming conventions are used to ensure efficient retrieval and evaluation of files.

4.3 Workflow

The user interacts with the system through a series of quiz levels. The workflow is as follows:

❖   The user logs in or registers, then begins the quiz from Level 1 (MCQs).

❖   After completing each level, the system evaluates the responses using GPT-3.5 and Whisper.

❖   The user advances to subsequent levels based on their performance and quiz completion.

❖   Upon finishing all levels, the system generates a certificate of completion if the user meets the required score threshold.

4.4 Security and Performance

To ensure security, user data and quiz responses are stored securely in the MySQL database with encryption. The system is optimized for performance, with efficient data retrieval from the database and minimal latency in AI model interactions.

5. Results

The login interface allows users to securely sign in or register to access the quiz system

The user dashboard provides access to Carnatic and Hindustani levels, games hub, fact generator
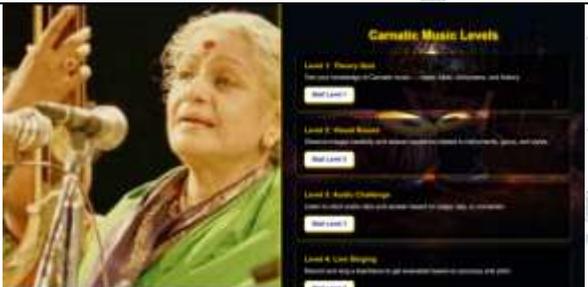
This interface presents Carnatic quiz levels

This interface presents Hindustani quiz levels
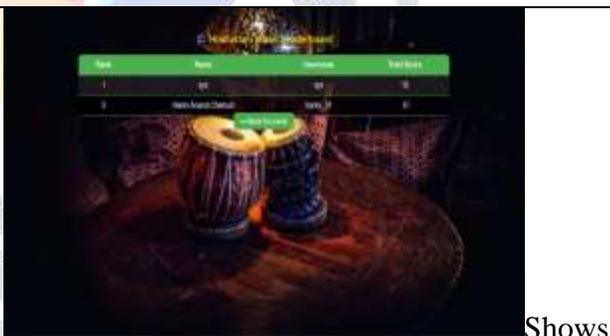
Displays top scores of users based on Hindustani music level performance

Shows AI-evaluated score and feedback after live singing in Level 4

6. Conclusion

The proposed system successfully integrates generative AI into Indian classical music education through a comprehensive four-level evaluation platform. By combining the powerful GPT and Whisper APIs with Java-based web development, the system facilitates interactive assessments across multiple modalities—text, image, audio, and live singing inputs. This enables personalized and accurate evaluation of user responses, including sophisticated pitch-wise analysis for sung ragas.

The system enhances traditional learning methods by offering a more accessible and engaging approach to music education. It not only modernizes the teaching process but also provides a scalable solution for evaluating knowledge in Indian classical music.

Looking ahead, future enhancements such as mobile support and advanced AI-driven feedback could further elevate the user experience, broadening the system's reach and impact. With these potential improvements, the system has the capacity to revolutionize how Indian classical music is learned, practiced, and assessed in the digital era.

7. References

Books and Publications

1.   **Bhatkhande, Vishnu Narayan.** *Hindustani Sangeet Paddhati*. Sangeet Karyalaya, 1910. A seminal work on Hindustani music theory, detailing the foundational concepts of ragas and their classification.
2.   **Sambamoorthy, P.** *South Indian Music – Volumes I–VI*. The Indian Music Publishing House, 1958. An extensive series that delves into the theory and practice of Carnatic music, offering insights into its rich traditions.
3.   **Raghavan, V.** *The Music of South India*. The Indian Music Publishing House, 1975. Explores the structure of Carnatic compositions and the intricate grammar of ragas and talas.
4.   **Arnold, Alison.** *The Garland Encyclopedia of World Music: South Asia – The Indian Subcontinent*. Routledge, 1999. A comprehensive reference that provides scholarly insights into South Asian musical traditions.
5.   **Saxena, Sushil Kumar.** *Hindustani Music and Aesthetic Theory*. D.K. Printworld, 2002. A modern philosophical study focusing on the aesthetics and performance theory of Hindustani music.

Technical Tools / Documentation References

1.   **OpenAI Whisper Documentation.**
Official documentation and usage guide for the Whisper speech-to-text model.

2.   **TarsosDSP Documentation.**
Java library used for audio pitch extraction and analysis.