

AIMC 2023

Introductory Studies on Raga Multi-track Music Generation of Indian classical music using AI

Sreekanth Gopi¹ Femi William¹

¹Kennesaw State University

Published on: Sep 09, 2023

URL: <https://aimc2023.pubpub.org/pub/4pcwrxuo>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Introductory Studies on Raga Multi-track Music Generation of Indian classical music using AI

Abstract

Recently, there has been an exponential expansion in research focusing on AI-based music generation. Our in-depth analysis of the arXiv dataset revealed a growing number of publications on this subject: over 273 AI music papers in the past two years, with 102 explicitly tackling AI music generation. However, one area that seems underrepresented is Indian traditional music. This study presents the application of artificial intelligence (AI) in creating Indian classical music, focusing on Raga-based music generation. We outline the two-stage music creation process, including the creative and technical aspects, and explore how AI can be integrated into these stages. We trained the models using the LSTM (Long Short-term Memory network) and Transformer models on the Dunya dataset, which includes almost 250 ragas played across 12 instruments. Further, the study proposes a new Raga Multi-Track Music Model (RMMM) model to generate multi-layered Raga-based music with enhanced authenticity and emotional resonance. Despite potential challenges, this research opens an exciting journey in AI-generated Indian classical music.

Keywords: AI Music, Music generation, Artificial Intelligence, Indian Traditional music, Carnatic Music, LSTM, Transformers, MMM, RMMM.

Introduction

In recent years generative AI(Artificial Intelligence) has received lots of attention, and its application to music generation through deep learning is gaining momentum. From an in-depth data analysis of the arXiv dataset of 2.25 million past STEM papers, we found that over 273 AI music papers in the past two years have been published, with 102 explicitly tackling AI music generation. However, our research found no instances of AI models being applied to composing Raga-based music. Raga, the heart of Indian classical music, creates unique emotional landscapes through a melodic framework consisting of specific note sequences, or 'Sarali Varasais' (Travis, Morehead, and Parim, 2023; Indira, 2018). Each raga, like Hindol and Todi, induces distinct emotions, including joy, playfulness, romance, and devotion (Sarkar and Biswas, 2015). We intend to explore these rich cultural expressions with the help of recent machine-learning techniques and make them easily available for global listeners and musicians.

A song's complex structure incorporates various layers, such as melody, rhythm patterns, and instrumentations, harmonized through orchestration (Xu, 2023) (Figure 1). The authors articulate music composition as a two-step process: creative exploration and technical production (Tables 1 and 2). Creatively, this process navigates through raga exploration, theme selection, main melody creation, instrument choice, and individual melody

compositions. And technically, it entails DAW-based track setup, rhythm, melody, and chord progression recordings, and the final mixing and mastering (Skillshare, 2023).

Step	Description
Raga Exploration	Experimenting with various ragas as per mood
Theme Selection	Choosing a fitting raga and theme
Melody Creation	Generating a heartfelt motif or main melody
Instrument Selection	Identifying suitable accompanying instruments
Tune Composition	Crafting melodies for each selected instrument

Table 1: Creative Composition

Step	Description	
Track Addition	Setting up individual tracks in DAW	
Rhythm Composition	Creating and recording the rhythm section	
Record Melody	Composing pads and recording the main melody	
Chord Progression	Playing and recording chord progressions	
Instrument Recording	Capturing each instrument's part	
Mixing and Mastering	Fine-tuning and polishing the final piece	

Table 2: Technical Production

AI can assist in each step of the composition process. Firstly, based on the creative selection process, AI can generate example melodies and patterns to create diverse tunes within raga constraints. These AI-generated melodies can then be recorded using live instruments in a DAW[1]. AI can also contribute to further processes such as producing rhythm patterns and chord progressions, adding effects, and finally mixing and mastering the output. Beyond helping the composer's process, these tools could independently produce an instrumental track that aligns with a chosen raga's unique construction.

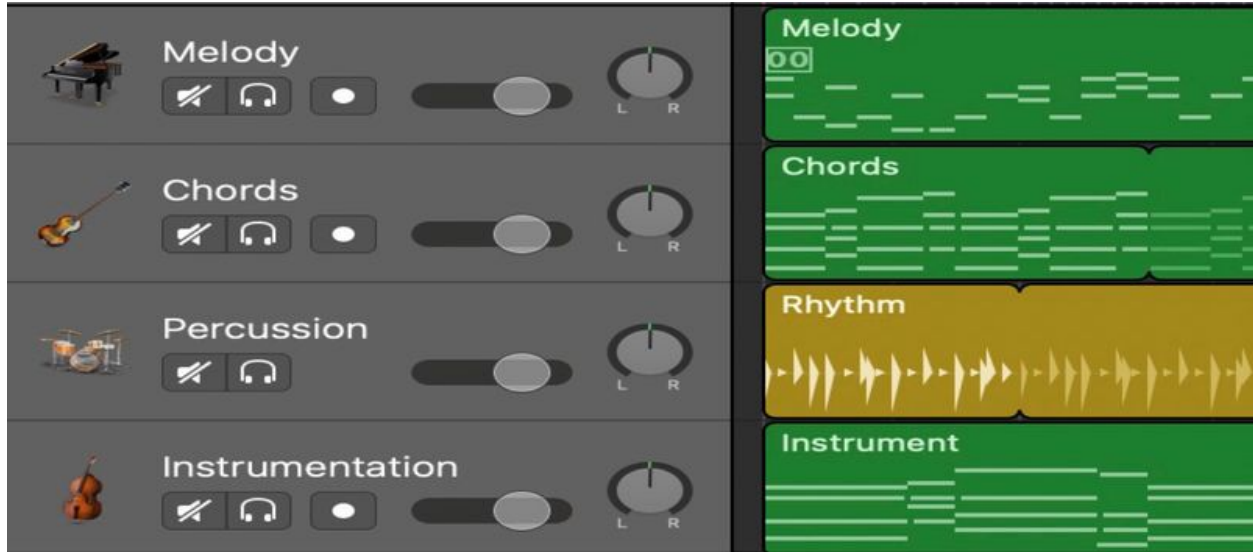


Figure 1: Song Structure example

This research marks a pioneering exploration into an underrepresented genre that could potentially expand current music generation models. Raga music is deep-rooted in ancient culture and possibly impacts the health and well-being of human beings (Sarkar, J., and U. Biswas., 2015). It is fading recently due to its obscurity, intricacy, and music modernization, revealing a need to develop tools to make it more accessible. Thus, it is imperative to develop ML models that can encapsulate the complexities of classical Indian music.

Methodologies

First, we examine AI music composition technologies with a rich multi-instrument Raga dataset. Then, we narrate LSTM and Transformer models researched, and their latest developments. Eventually, we propose a novel system, RMMM, specifically designed for raga-based multi-track music.

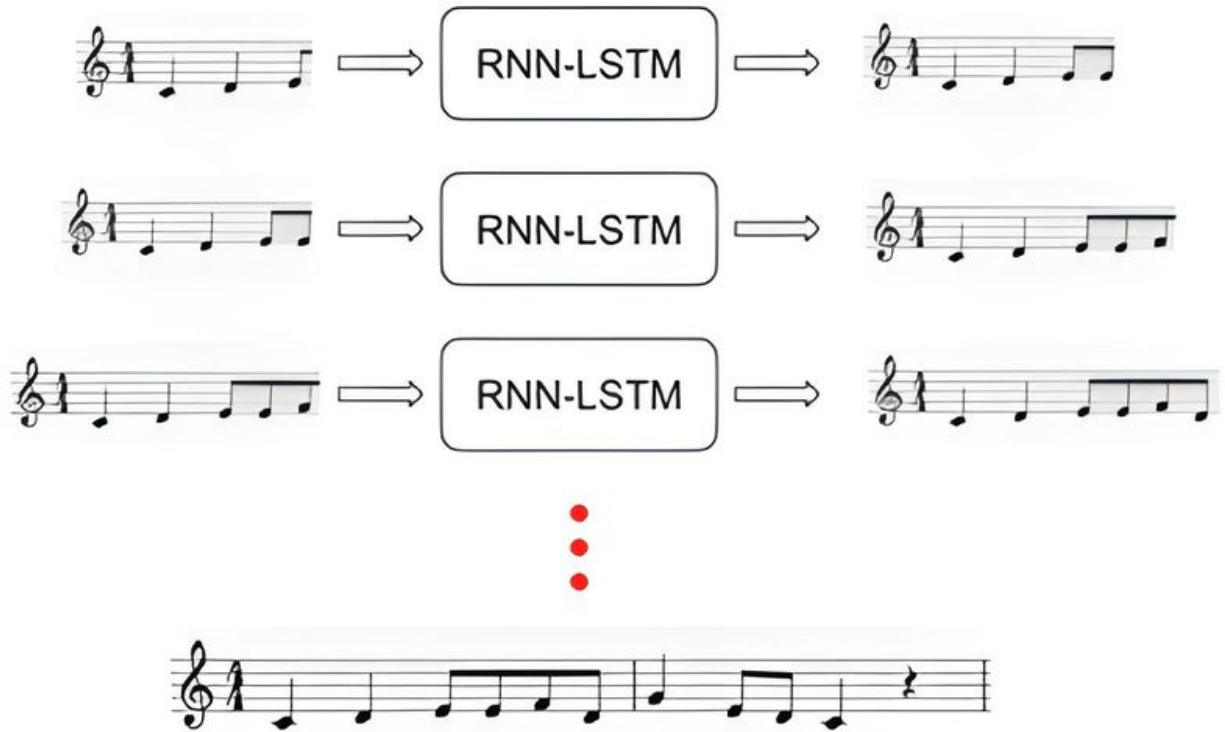
Dataset

Dunya, part of the CompMusic project, is a significant resource for studying Indian traditional music. It hosts almost 250 ragas played across 12 diverse instruments, including the flute, violin, tabla, etc. It is organized into its categories facilitating detailed musicological analysis. Its robust API allows seamless access to this vast collection, making it ideal for Music Information Retrieval tasks in raga research(CompMusic, 2023).

LSTM Music Generation Technology

Our analysis focused on the latest technologies of AI music composition using machine learning models. LSTM models are efficient at learning long-term dependencies and are employed for melody generation due to their ability to capture repetitive musical structures over time. The melody represented as a time series of MIDI notes and rests, is processed by the LSTM. Each step, representing a 16th note, is sequentially fed into the

model. The model predicts the following note based on the learned musical context, generating a sequence of note-by-note (Figure 2, Velarado, 2020). This iterative process enables the LSTM to produce a coherent melody, reflecting the long-term structural patterns inherent in music (Figure 4, Velarado, 2020).



LSTM melody generation

Figure 2: LSTM sequential note generation.

First, we explored an LSTM-based Performance-RNN[2] called the Google Magenta project, designed to model polyphonic music with expressive timing and dynamics. The system achieves this by generating a stream of MIDI events, including note-on and note-off events for each of the 128 MIDI pitches, time-shift events, and velocity events. The events are represented as one-hot vectors in a MIDI-like stream, allowing for the fine quantization of expressiveness in note timings. The model can compose performances directly, determining which notes to play, when, and with what intensity. The generated music (saved in our GitHub repo) produced with chords seems suitable for generating raga-based music as it deals with expressive timing and dynamics, often used in raga singing. Although the model can generate longer performances, they lack note raga renditions and long-term structure (Table 3) (Simon and Oore, 2017).

Feature	LSTM	MMM
Model Type	RNN	Transformer
Long-Term Dependencies	Yes	Yes
Sequences of Arbitrary Length	Yes	Yes
Aligns Songs to a Standard Key	Yes	Yes
Filters Out Tunes	Yes	Yes
Variability in Note Generation	Yes	Yes
Multi-Track Music	No	Yes
User Control	No	Yes

Table 3: LSTM vs MMM

Transformer Models

To explore the creation of multi-track music, we experimented with the paper (Ens and Pasquier, 2020) that introduces MMM, a generative system based on the Transformer architecture (Figure 3). It crafts individual time-ordered sequences for each track, that are woven together into a single harmonious polyphonic music. In our code execution, we found that MMM can be used to generate harmonious Raga music by training the system on specific Ragas. The MMM4Live UI and Calliope version (Tchemeube, Ens, and Pasquier, 2022) were found to produce raga-based tracks[3] based on a given input. This has the potential to generate different instruments based on a selected range of polyphony, density, note length, temperature, etc (sample songs are saved at:<https://github.com/datasci888/AIMusic>).

Novel RMMM Model

To venture into the realm of Raga-based multi-track music, we propose the development of the Raga Multi-Track Music Machine (RMMM) as shown in Figure 3, an extension of the Multi-Track Music Machine (MMM) architecture(Ens and Pasquier, 2020). Building on MMM’s unique feature, the RMMM would skillfully interlace separate, temporal sequences from each track, culminating in a harmonious, polyphonic

Raga composition. This system would be trained on datasets of Raga-based music and fine-tuned based on feedback from Raga experts.

An intuitive user interface (UI), RMMMLive, would allow users to generate Raga-based multi-track music and provide control over iterative re-sampling, to be developed based on MMM4Live and Calliope (Tchemeube, Ens, and Pasquier, 2020). The interface would support the creation of music aligning with specific moods, and purpose-based song creation. Its enhancements would include auto-improvisation, integration of rhythm, microtonal adjustment[\[4\]](#), and cross-cultural fusion features.

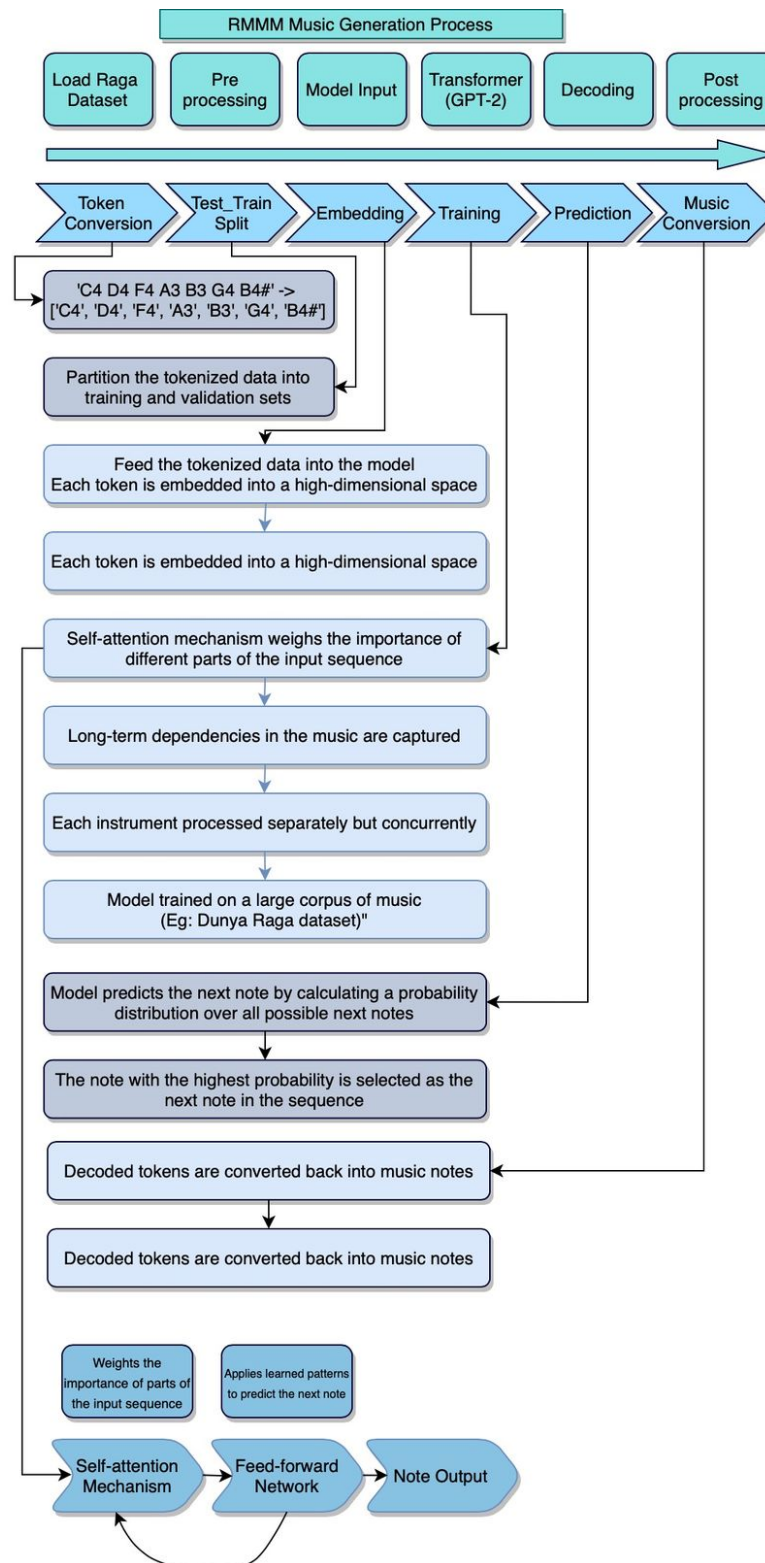


Figure 3: RMMM Architecture

Tool	Features
EmotionBox	Analyzes musical elements to determine and compose emotional characteristics.
MMM	Creates multi-track music with user control over various aspects.
MSAT	Enhances music generation quality by balancing harmony and coherence.
MMT	Generate longer multi-track music at a faster inference speed
MMM4Live	Interactive platform for real-time generation and manipulation of music.
Calliope	Web application for multi-track music composition with MMM.

Table 4: ML Models

Discussion

MMM is better suited for generating Raga-based music than LSTM due to its ability to harmonize polyphonic music and tailor specific parameters (Table 3). Expanding upon MMM with features as shown in Table 4, we propose the RMMM architecture (Figure 3) that would introduce innovative automation in the Raga-based music composition process. Additionally, LSTM PerformanceRNN (Simon and Oore, 2017) could also be integrated into RMMM for simpler melody generation in cases of simple melodies, smaller datasets, and lower computational requirements. Integration of Emotionbox, a system that uses an improved RNN to generate emotion-specific music (Zhenget. al., 2022), could enhance emotion selection. Even more, integration with Multi-Scale Attentive Transformer (MSAT) technology (Wei. al., 2023), and Multitrack Music Transformer (MMT) (Dong et al., 2023) could further enhance the performance of the RMMM. Besides, extensive research could be done for developing RMMM to accept types of input such as scores, lyrics, text, images, and audio. These integrations could be useful to create more layered, long, and complex Raga compositions.

Conclusion

The eclectic Raga system offers thousands of musical scales, each tied to specific melodies, moods, meditation, healing, devotion, sleep, and even specific hours of the day. A listener or a composer seeking new music could use RMMM to create it even with limited knowledge of Ragas. Eventually, we plan to publish pre-trained models specific to each raga, providing users with quick and easy usage for different applications including

Raga music therapy. Performers could use RMMM for creating an instrumental track or a karaoke to add an innovative, AI-driven dimension to their live shows.

References

- Weiet. et al., 2023. “A Multi-Scale Attentive Transformer for Multi-Instrument Symbolic Music Generation.” *Journal of Artificial Intelligence Research*.
- Zhenget. et al., 2022. “EmotionBox: A Music-Element-Driven Emotional Music Generation System Based on Music Psychology.” *Frontiers in Psychology*.
- Dong, Hao-Wen, Ke Chen, Shlomo Dubnov, Julian McAuley, and Taylor Berg-Kirkpatrick. 2023. “Multitrack Music Transformer.” In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- CompMusic. (n.d.). Dunya. Retrieved June 29, 2023, from <https://dunya.compmusic.upf.edu>
- Ens, Jeff, and Philippe Pasquier. 2020. “MMM: Exploring Conditional Multi-Track Music Generation with the Transformer.” *arXiv Preprint arXiv:2008.06048*.
- Indira, A. et al. 2018. “Effectiveness of Music Therapy on Academic Performance of Nursing Students.” *International Journal of Academic Medicine*.
- Sarkar, J., and U. Biswas. 2015. “Indian Classical Ragas to Cure Diseases.” *Int. J. Adv. Sci. Res* 1 (1): 9–13.
- Simon, Ian, and Sageev Oore. 2017. “Performance RNN: Generating Music with Expressive Timing and Dynamics.” *Magenta Blog*. <https://magenta.tensorflow.org/performance-rnn>.
- Skillshare. 2023. <https://www.skillshare.com/en/blog/how-to-compose-music-a-step-by-step-guide/> \newline (accessed:06.27.2023).
- Tchemeube, R., Jeffrey John Ens, and P. Pasquier. 2022. “Calliope: A Co-Creative Interface for Multi-Track Music Generation.” In *Proceedings of the 2022 ACM SIGCHI Conference on Creativity and Cognition*. ACM.
- Travis, F., P. Morehead, and N. Parim. 2023. “Effects of Gandharva Veda Music on Mood States, Health, and Brain Functioning.” *International Journal of Psychological Studies*.
- Velarado, Valerio. 2020. “Generating Melodies with RNN LSTM.” *GitHub Repository*. <https://github.com/musikalkemist>; GitHub.
- Xu, Yurui. 2023. “Music Generator Applying Markov Chain and Lagrange Interpolation.” *HSET*. <https://doi.org/10.54097/hset.v39i.6538>.

LSTM Architecture

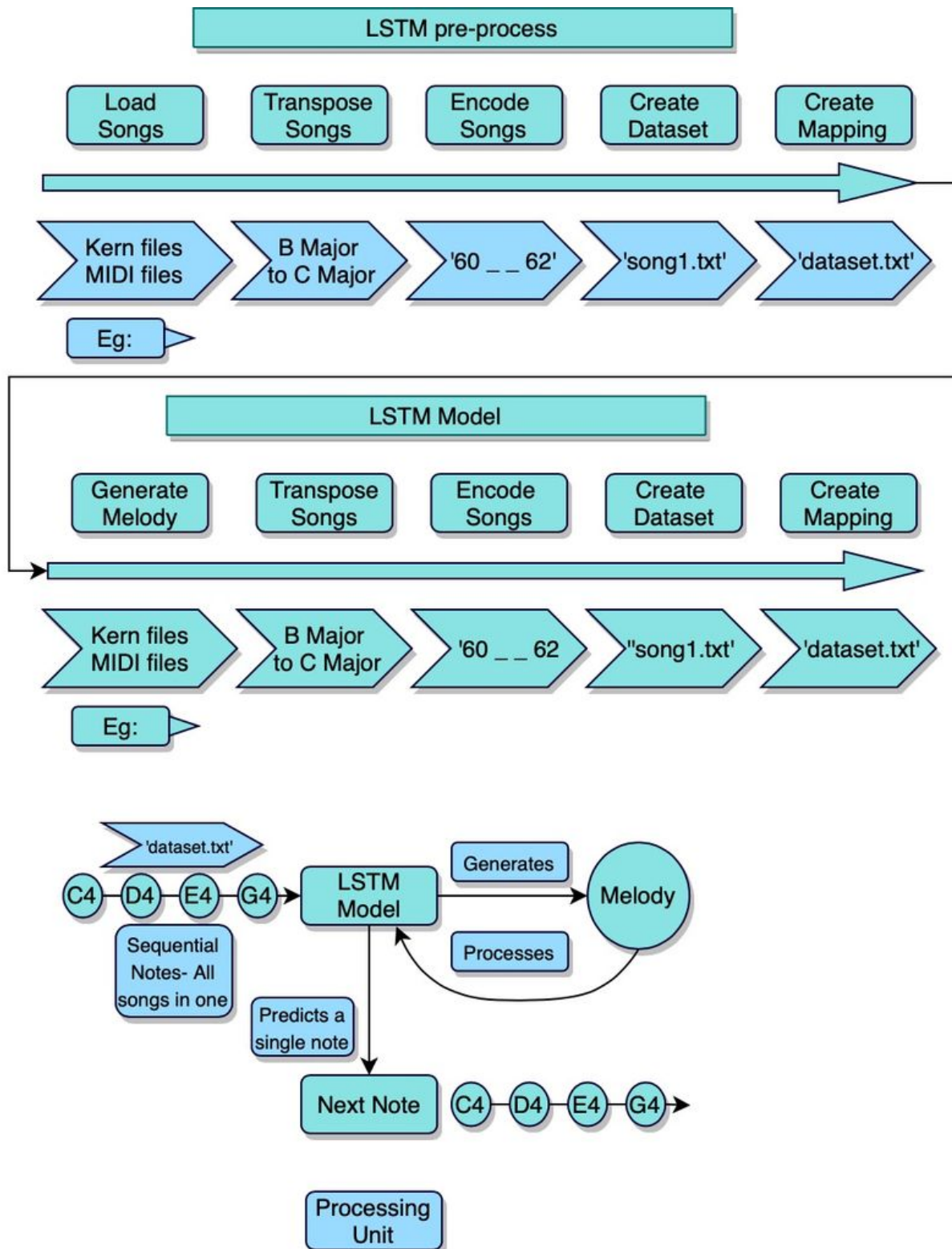


Figure: 4

Footnotes

[1] Digital Audio Workstation

[2] Recurrent Neural Network

[3] The Generated music tracks are available at: <https://github.com/datasci888/AIMusic>

[4] Microtonal Adjustment: This feature enables the usage of microtones (shrutis) in compositions, fostering greater authenticity and nuance.
