

Low-Resource Rhythm Learning: Machine Learning Approaches to Structured Classical Beats

Deriving semantic representations of rhythmic structures is essential for AI-driven choreography and music generation. In South Asian classical dance, rhythms are not only an accompaniment but the key element of choreography, providing the temporal scaffold on which sequences are composed and improvised. Building computational models of these rhythms is thus critical for devising AI agents that can generate, complete, or interact with dance. Yet, unlike Western rhythmic corpora that are well-studied, South Asian traditions remain underexplored in machine learning, largely due to the absence of annotated datasets and standardized representations.

We focus on *nattuvangam*, the vocal percussion and instrumental cues that conduct Bharatanatyam dance, as a concrete instantiation of structured beats. Nattuvangam cycles are articulated via syllables (bols) and percussion, coordinating dancers in real time and shaping expressive phrasing. Similar principles exist in Kuchipudi, Mohiniyattam, and Kathak, indicating a generalizable framework for rhythm modeling beyond any single dance form. However, existing rhythm learning approaches, designed for large symbolic datasets, fail to capture cyclical, culturally specific patterns under extreme data scarcity.

To address this gap, we construct the first curated dataset of nattuvangam recitations, applying audio cleaning, segmentation, and transcription pipelines to extract temporally aligned sequences. Leveraging this resource, we train autoregressive transformer-based models with combined audio and symbolic embeddings to learn cyclical rhythmic progression. The models are optimized using a joint objective-cross-entropy loss over symbolic tokens and reconstruction loss over audio embeddings. Evaluation with cycle prediction accuracy and generative fidelity shows that incorporating structural priors on cyclical beats improves prediction and generalization under low-resource conditions.

This work complements efforts such as *Footwork2Framework*, which captures 13 hours of fine-grained 3D Bharatanatyam motion data, by modeling the auditory conductor that guides movement. We aim to collect a comparable scale of synchronized audio to extend our framework, creating a holistic computational representation of both rhythm and motion. More broadly, this methodology generalizes to other classical music and dance traditions where cyclical vocal or instrumental cues structure performance, enabling rhythm learning in domains with little to no existing data.

By bridging culturally specific rhythmic knowledge with machine learning, this research advances human-AI co-creation and choreography generation, while also contributing to the preservation of intangible cultural heritage.