Audix: AI-Enabled Music Platform with Music-to-Symbolic Notation Conversion

Ashna Jacob
PG Scholar
Department of Computer Application
Amal Jyothi College of Engineering Autonomous
Kanjirappally, Kerala, India
ashnajacob2026@mca.ajce.in

Dr. Paulin Paul
Assistant Professor
Department of Computer Application
Amal Jyothi College of Engineering Autonomous
Kanjirappally, Kerala, India
paulinepaul@amaljyothi.ac.in

Abstract*— Music streaming services are increasingly the most interactive media for contemporary entertainment and education. Yet the majority of services emphasize content delivery and personalization, neglecting the technical and learning sides like music notation generation for students and professionals alike. Audix bridges this gap with an embedded Music-to-Symbolic Notation conversion system that translates audio to symbolic sheet music. Besides this, Audix features intelligent modules like Voice Match for complete song recognition from partial audio clips, AI-driven recommendation, real-time conversation, playlist control, and social engagement. The system facilitates learning, discovery, and engagement on a single platform. Architecture is of a three-tier type with presentation, application, and data layers. Core AI technologies—deep neural networks, symbolic music processing models, and recommendation algorithms—facilitate accurate notation conversion and interactive user experience.

***Keywords*—AI, Music-to-Notation, Symbolic Music, Voice Match, Music Recommendation, Real-Time Chat**

I. INTRODUCTION

The digital music market has changed the way people access and interact with music. Services such as Spotify, Apple Music, and YouTube Music are focused on discovery by individual tastes, skilled playlists, and social sharing. These services barely accommodate notation-based learning tools, vital for beginners, learners, and professionals in the music education arena.

Audix fills this gap by the incorporation of a music-to-symbolic notation transcription system that translates any input audio into readable symbolic sheet music. This functionality enables users to study melodies, examine rhythm, and practice professionally. Moreover, the Voice Match technology enables users to recognize a full track from even a small audio input or humming, thus simplifying discovery and making it enjoyable.

Audix is built as an interactive platform to serve user management, subscription, AI recommendation, music library management, playlist management, and community chat modules. With artificial intelligence and symbol representation methods, Audix enables amateur listeners and professional learners to engage and benefit from its ecosystem.

II. LITERATURE REVIEW

Recent developments in symbolic music creation with AI mirror its capability for learning and information extraction. Symbolic representation is seen to offer a structured and language-like syntax for music that allows systems to transcribe and create musical scores from raw audio inputs.

Systems like Google's "Hum-to-Search" illustrate strong neural design with the capability of transcribing hummed music or hummed sounds into recorded music. Likewise, symbolic music research focuses on sophisticated deep neural networks for rhythm segmentation, pitch segmentation, and note onset detection.

Toiviainen's survey offers symbolic AI solutions to segmentation, classification, and recommendation problems. Clever music discovery interfaces like Cyanite.ai and ISMIR models demonstrate practical uses of music content annotation and notation in real life.

Audix differs from other music streaming applications by combining symbolic notation with social community interaction for developing an ecosystem where formal learning is accompanied by personalization. The two-sided nature of the approach enables learners, producers, and occasional listeners to profit at the same time.

III. METHODOLOGY

Audix system is based on three-tier architecture yielding modularity, reliability, and scalability. Presentation tier offers music streaming interface foplayback, notation viewer panel for displaying music converted, playlist controls like create, reorder, like, and share, and real-time chat support as well as secure user authentication interface. The application layer collects the essential features such as deep learning-based Music-to-Symbolic Notation engine for audio to symbolic note conversion, Voice Match algorithm for song matching with partial audio inputs and full songs, hybrid and content-based filtering algorithms-powered smart recommendation engine, and playlist plus library management modules for smooth user experience. The data tier enables scalability and storage through cloud storage repositories for symbolic notation files and annotated datasets, a relational database for storing user accounts, session usage, and preferences, and scalable media storage for music collections and audio datasets.

The platform is powered by state-of-the-art technologies like HTML5, CSS, JavaScript, and ReactJS in the frontend for dynamic rendering, Python Flask or Node.js with Firebase for the backend, and sophisticated TensorFlow/Keras deep learning models for the processing of symbolic notation. There are also other AI/ML modules that are responsible for note detection, music recognition, and symbolic sequence representation based on NLP models, and the overall security framework provides multi-factor authentication, secure API communications, and encryption of sensitive user data.

A. Data Collection

The Audix system gathers data from large annotated sets of music such as MAESTRO and MusicNet, which include music tracks and symbolic notation files corresponding to them. The tracks uploaded by the users are also employed for incremental model training.

B. Data Preprocessing

The audio signals are processed by preprocessing methods like noise removal, filtering, and adjustment of time-stretch. Pitch contours, note onset, and chroma information are the features that are extracted. The features are normalized to keep the genres consistent, and annotated scores are aligned with the audio samples carefully.

C. MUSIC-TO-SYMBOLIC NOTATION PROCESS

The music-to-symbolic notation process in Audix includes some major steps:

Audio Input Capture: Audio files or live recordings are fed into the system.

Feature Extraction: Pitch detection algorithms extract notes independently.

Segmentation: Onsets and offsets are detected to determine rhythm structure.

Mapping: Symbolic units like notes, rests, and tempo are mapped from detected musical features.

Score Creation: The symbolic data are exported as standard musical notation for display or download.

It is checked for accuracy against dataset references. More complex AI models—such as VGGish embeddings, RNNs, and Transformers—are used to further develop note detection and rhythm accuracy.

IV. EVALUATION

Audix gathers data from enormous annotated music databases like MAESTRO and MusicNet, which consist of music tracks synchronized with symbolic notation files, and again trains itself using user-supplied tracks for ongoing incremental learning. The audio signals are subjected to several preprocessing phases of noise reduction, filtering, and time-stretch adjustment prior to analysis in order to maintain clarity and consistency. The discriminative features like pitch contours, onset time, and chroma are derived from analyzed signals, and the features are normalized to make them comparable across music genres. Proper alignment of annotated symbolic scores and corresponding audio samples allows for proper setting of proper relationships between recorded music and symbolic representation.

The Audix process of symbolic notation-to-music starts with the input audio being captured either as a file or live capture. The input is then subjected to feature extraction, in which the system extracts notes from the input through sophisticated pitch detection algorithms. Detected onsets and offsets segment the music, which are central to organizing rhythm. Every detected musical item is projected onto symbolic units including note, rest, and tempo markings. Lastly, the system generates a final musical score by printing out symbolic data as common notation understandable by musicians. The resulting symbolic output is checked against annotated dataset references to ensure that it is correct, and sophisticated machine learning models such as VGGish embeddings, RNN-based transcribers, and Transformer networks improve note detection precision and rhythm stability throughout the process.

V.RESULT

The deployment exhibited great performance in both symbol notation and recognition tasks. The 83% notation correctness demonstrates good transcription capability, and a latency of less than 1 second demonstrates real-time usability. User surveys also validated that students found automatically generated sheet music to be highly useful for practicing and learning exercises.

VI.CONCLUSION

The proposed Audix system demonstrates the potential of merging music streaming with professional learning tools. By integrating music-to-symbolic notation conversion, intelligent recommendations, and social features, the system provides both entertainment and educational benefits.

The evaluation shows promising results in accuracy, response time, and usability. This makes Audix an ideal next-generation platform suitable for learners, performers, and general users. Future scopes include extending symbolic support to polyphonic orchestration, real-time MIDI-to-notation rendering, and integration with AR/VR environments for immersive training.

REFERENCES

[1] Symbolic Music Generation Using Deep Neural Networks. TowardsAI. 2024.

[2] Restackio. Symbolic Music Generation with AI.

[3] Ting Zhang. An innovative new AI tool that recognizes musical notation. International Journal of Wireless and Mobile Computing. 2024.

[4] Google AI Blog. Song stuck in your head? Just hum to search. 2020.

[5] Petri Toiviainen. Symbolic AI Versus Connectionism in Music Research. Dept. of Musicology, University of Jyväskylä.