National Conference on Advances in Science, Technology, Engineering, and Management

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

MELODY-MIXER: Music Generation Using Deep Learning

Sriraksha, Sahana H S, Rakshitha R, Vanditha, Amitha S*

(Department of Computer Science and Engineering, Kammavari Sangham School of Engineering and Management, Bengaluru, India

Email: srirakshashetty57@gmail.com, hssahana316@gmail.com,rakshitharajesh02@gmail.com, vandithamogaveera@gmail.com, amitha.s@kssem.edu.in*)

****

Abstract:

Music generation using Deep Learning presents a comprehensive system for generating music in ABC notation using character based recurrent neural networks (RNNs), accompanied by a conversion pipeline that transforms the output music from ABC notation to a playable audio format. The integration of character-based RNNs allows for the creation of coherent and melodic musical compositions, while the ABC-to-MIDI-to WAV conversion enhances the usability and accessibility of the generated music. The work starts by preprocessing the ABC notation dataset and representing it at the character level. A character-based RNN, such as a long short-term memory (LSTM) network, is then employed to learn the sequential dependencies within the ABC notation and generate music that follows the learned patterns and structures. The RNN model is trained on a substantial corpus of ABC encoded music, enabling it to capture the statistical regularities and nuances of the dataset. To make the generated music readily playable, the project incorporates a conversion pipeline that translates the output from ABC notation to MIDI format. MIDI files serve as a widely supported industry-standard representation of music, making them compatible with a variety of digital audio workstations (DAWs) and synthesizers. The MIDI files are subsequently converted to WAV format, a universally recognized audio format suitable for playback on diverse platforms and devices. This project contributes to the field of AI-generated music by providing an integrated system for music generation in ABC notation, accompanied by a seamless conversion pipeline for playback in the universally supported WAV audio format.

Keywords — **Music Generation, Python, Streamlit, Recurrent Neural network (RNN), Long-short Term memory (LSTM).**

****

I. INTRODUCTION

Music has played a fundamental role in cultural expression and societal development. While the idea of machines generating music through algorithms once seemed fantastical, it has gradually transitioned from a distant dream to a tangible reality. Over the past century, countless theories, papers, and implementations have emerged in this unique field. Despite imperfections, significant strides have been made in the last decade, leading to the emergence of various models and applications capable of generating music. However, challenges such as excessive noise and extensive processing time persist. The project endeavors to design and train a character-based RNN model capable of learning the intricate patterns and structures inherent in ABC notation. Leveraging a dataset sourced from the Nottingham Music Dataset, which encompasses a vast collection of 1200 folk tunes in ABC notation, the RNN model aims to encapsulate the core characteristics of various musical genres. The RNN model architecture will utilize the Keras sequential model, a flexible deep learning framework.

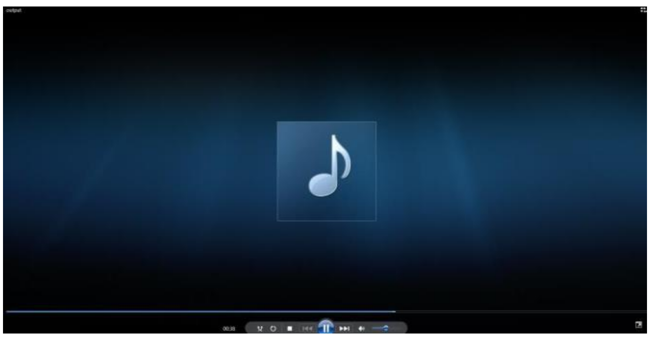


Fig 1:Music sequence.

LITERATURE SURVEY

Deep Learning (DL), a sub-field of Machine Learning (ML), has been established as a strong computational toolbox, with applications in numerous tasks, like feature extraction, classification, and pattern recognition. Such functionalities enable the extraction of meaningful information from raw data, and thus find applications in a wide range of disciplines, including computer vision (CV), natural language processing (NLP), bioinformatics, medical diagnosis, speech recognition, image processing (IP), system identification, recommendation systems, and more. A research field where DL has emerged as a valuable tool over the last decade is that of audio signal processing (ASP) and music signal processing (MSP). Music is a well-known art form that is a big part of the most fun and educational human activities. As a result, the music industry includes a wide range of organizations and consumers. The application of DL tools in MSP has led to a collection of successful commercial applications, the most famous of which is Music Recommendation Systems (MRS)

I. TECHNOLOGY STACK

Python- Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high- level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

ABC Notation - ABC notation is a shorthand form of musical notation for computers. In basic form it uses the letter notation with a–g, A–G, and z, to represent the corresponding notes and rests, with other elements used to place added value on these – sharp, flat, raised or lowered octave, the note length, key, and ornamentation. This form of notation began from a combination of Helmholtz pitch notation and using ASCII characters to imitate standardmusical notation.

MIDI - ABC notation is a shorthand form of musical notation for computers. In basic form it uses the letter notation with a–g, A–G, and z, to represent the corresponding notes and rests, with other elements used to place added value on these – sharp, flat, raised or lowered octave, the note length, key, and ornamentation. This form of notation began from a combination of Helmholtz pitch notation and using ASCII characters to imitate standardmusical notation.

WAV - WAV files are the standard audio format for high- quality, uncompressed sound on Windows PCs. Originally created by IBM and Microsoft in 1991, the.wav file extension stands for " waveform audio file," referring to thewaveform representation of the audio data.

II. SYSTEM ARCHITECTURE

The overall workflow can be divided into these main steps:

- **Model**: The Model is loaded and initialized with the trained Datasets. It accepts the initial character entered by the user to be passed into the evaluator function for prediction of further sequence of characters.
  The sequence length basically denotes the length of thepredicted music sequence.

- **Output in ABC-Notation**: ABC Notation serves as a condensed form of musical notation designed for computer use. At its core, it employs letters a–g, A–G, and z to denote notes and rests, while additional elements like sharps, flats, octave adjustments, note duration, key, and embellishments enhance its expressiveness. Originating from a fusion of Helmholtz pitch notation and ASCII characters mimicking traditional musical notation, ABC facilitates online music sharing and offers a straightforward language for software developers, akin to other user-friendly notations.
- **MIDI Output Format**: MIDI, an abbreviation for Musical Instruments Digital Interface, serves as a standard for transmitting and storing music, initially devised for digital music synthesizers. Following the loading of the generated ABC notation output, we utilize a library known as

music21. This library facilitates the conversion of ABC notation into a playable MIDI format, allowing the music to the format

- WAV Output format: WAV Output format is basically considered for storing acoustic waves, and the name is pronounced" wave." The file extension is. wav. It's all simple and logical. We cannot pass the MIDI file directly as input to the web application because by nature of most web application MIDI sound format or any other sound formats other than WAV is not supported.
- Playable Audio: The generated WAV file is transmitted to the developed web application, providing users with a suitable interface to specify their preferred instrument, pitch, and desired sequence. These elements are crucial components of the music, which users can preview and download according to their preferences.

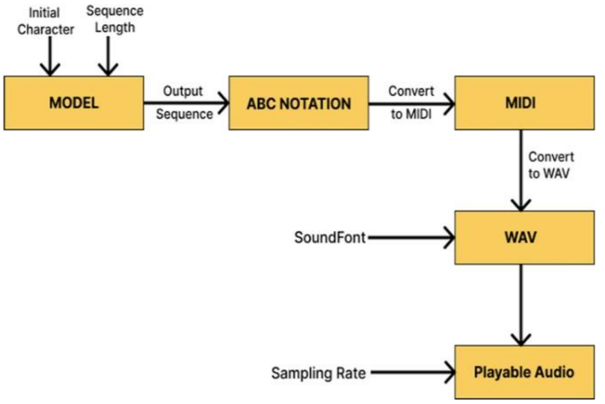


Fig 2: System Architecture for Music Generation.

III. PROPOSED SYSTEM

Music generation through deep learning initiates with user interaction, where users input preferences such as genre or desired musical elements. This input undergoes preprocessing to prepare it for the deep learning model, involving tasks like parsing textual descriptions or encoding categorical variables. At the heart of the system lies the deep learning model. Following music generation, sequences undergo postprocessing to refine quality or ensure compliance with musical conventions, such as smoothing note transitions or adding dynamic variations. The resulting music is presented to users for evaluation and feedback. Users listen, provide feedback, and may adjust input parameters to refine the generation process. Based on user feedback, the system iterates through the generation process, tweaking parameters or refining the model to enhance music quality. Once users are satisfied, the final output can be saved or presented in a suitable format for further use or distribution. Throughout this iterative journey, the system seamlessly integrates deep learning techniques to craft music tailored to user preferences and quality standards.

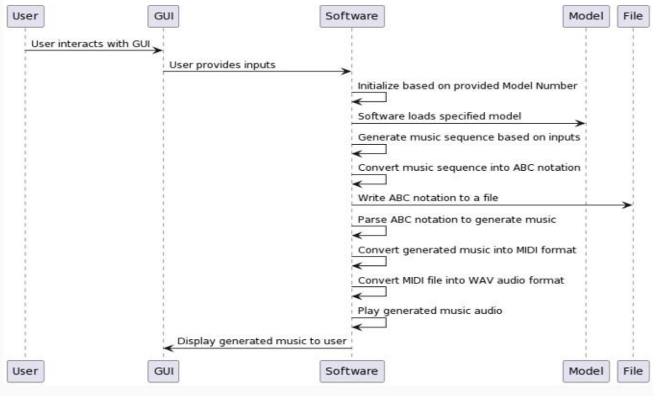
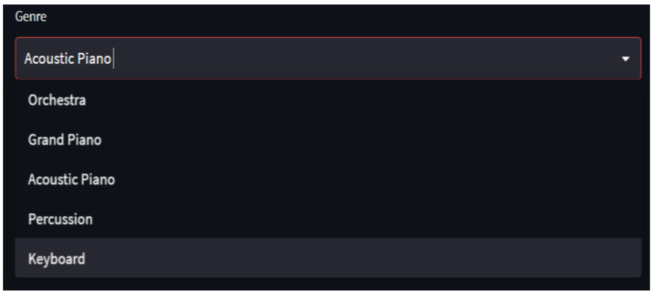


Fig 3: Sequence diagram.

IV. RESULTS AND SNAPSHOTS



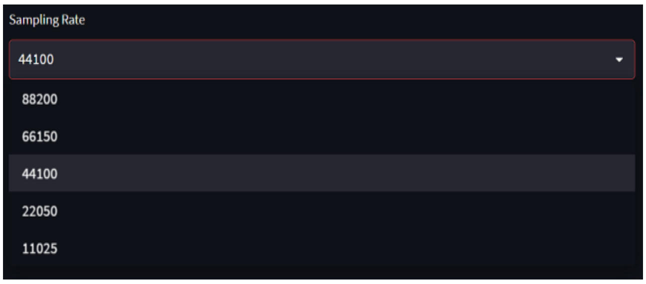


Fig 4.1: Different Genres Available to the User.

Fig 4.2: Different sampling rates available to the user.

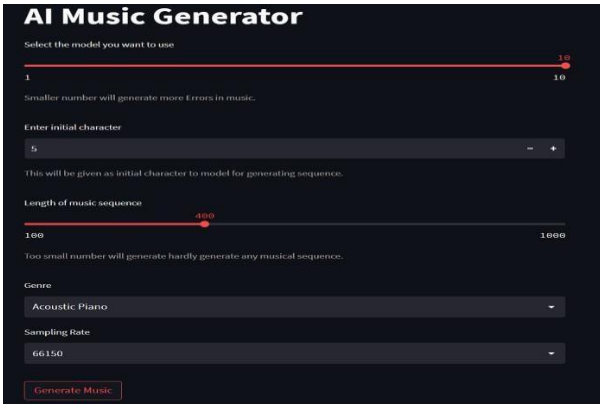


Fig 4.3: Music Generator Input given by the User.

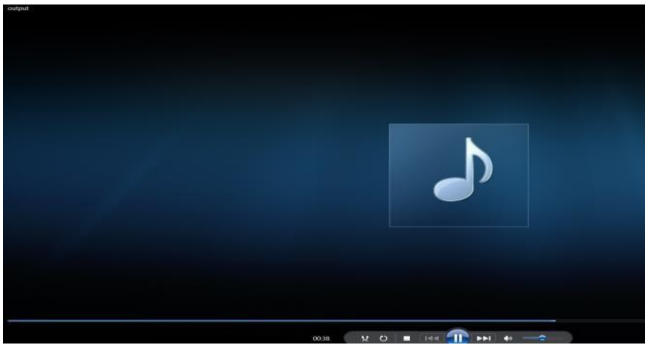


Fig 4.4: Playable Audio generated by Music Generator.

V. CONCLUSION AND FUTURE ENHANCEMENTS

From a social perspective, the project can be considered feasible as it aims to provide a creative and engaging experience for music enthusiasts. The generated music can be enjoyed by individuals seeking new folk tunes, musicians looking for inspiration, and music researchers interested in exploring the capabilities of AI in music composition. However, the social acceptance and adoption of AI- generated music may vary among different communities and individuals, and it is essential to consider the ethical implications and copyright considerations when generating and sharing music.

The application of deep learning techniques to music generation represents a groundbreaking approach with several significant contributions. Firstly, deep learning models have the capacity to learn intricate patterns and structures from large-scale musical datasets, enabling the generation of compositions that exhibit a high degree of complexity and creativity. These models can capture subtle nuances in musical styles and genres, allowing for the creation of diverse and authentic pieces of music. Additionally, deep learning-based music generation systems facilitate the exploration and fusion of different musical elements, paving the way for innovative compositions that transcend traditional boundaries. Moreover, these systems serve as valuable tools for both professional musicians and enthusiasts, offering a platform for experimentation, inspiration, and collaboration. By democratizing the process of music composition, deep learning contributes to the evolution of musical expression and fosters new forms of artistic creativity.

VI. REFERENCES

[1] Ji, Shulei, Jing Luo, and Xinyu Yang. “A comprehensive survey on deep music generation: Multilevel representations, algorithms, evaluations, and future directions.” arXiv preprint arXiv:2011.06801 (2020).

[2] Shahriar, Sakib. “GAN computers generate arts? a survey on visual arts, music, and literary text generation using a generative adversarial network.” Displays 73 (2022): 102237.

[3] Hernandez-Olivan, Carlos, Javier Hernandez-Olivan, and Jose R. Beltran. “A survey on artificial intelligence for music generation: Agents, Domains and Perspective.” arXiv preprint arXiv:2010.13944 (2022).

[4] Wen, Yu-Wei, and Chaun-Kang Ting. “Recent Advances of Computational Techniques for Composing Music.” IEEE Transactions on Emerging Topics in Computational Intelligence (2022).

[5] Moysis, Lazaros, Lazaros Alexios Iliadis, Sotirios P. Sotiroudis, Achilles D. Boursianis, Maria S. Papadopoulou, Konstantinos-Iraklis D. Kokkinidis, Christos Volos, Panagiotis Sarigiannidis, Spiridon Nikolaids, and Sortirios K. Goudos. “Music Deep Learning: Deep Learning Methods for Music Signal Processing- A Review of the State-of-the- Art.” IEEE Access (2023).

