IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Indian Classical Music Generation Using LSTM and RNN

BHAVANASREE, KRISHA, SHANAZ, KRISHNAN

STUDENT, STUDENT, STUDENT, STUDENT

COMPUTER ENGINEERING DEPARTMENT,

Presidency University, Bengaluru, Karnataka

Abstract: In the contemporary music landscape, it's a fallacy to believe that only those with expert knowledge in music can create high-quality tunes. Music enthusiasts, regardless of their expertise, can produce music that resonates. Music is a universal language that many enjoy, and the advent of automatic music generation stands as a potential revolutionary milestone in the music industry.

Traditionally, music was created manually using analog processes. However, recent advancements have seen music production pivot towards digital methodologies, with technology playing a supporting role alongside human creativity. This paper outlines the development of generative neural network frameworks capable of capturing the intricate aspects of harmony and melody autonomously. It provides a concise overview of music's fundamental characteristics, citing relevant sources where necessary. The core aim of the research is to develop a method for generating musical compositions using Long Short-Term Memory (LSTM) networks within the framework of Recurrent Neural Networks (RNNs).

The generated compositions are initially in ABC notation, a method for music representation, which is subsequently transformed into MP3 format for broader accessibility. ABC notation serves as the chosen format for training the model, splitting into metadata that details the characteristics of the music, such as index, time signature, default note length, and tune type, and the actual musical notes represented by a series of characters. The algorithm is designed to master the patterns of monophonic music through a trio of single-layered LSTM networks, showcasing its ability to learn and replicate musical sequences efficiently.

I. INTRODUCTION

We aim to design a network that could automatically generate piano music. The network is trained based on a predictive task: we fit the network to predict the representation of the note to come based on the previous notes that have been played. We then use this predictive model to generate notes that are likely to appear in a human-composed piece. Furthermore, studying the relationships between time-dependent events in music, and learning how to find automatically patterns and structures in music tracks through Deep Learning approaches is probably useful in other domains like finance

.

A particularly interesting bridge between the two fields is that piano recording analysis has to deal with different time scales [1], to understand links between events happening at the scale of a few milliseconds, to broad relationships on the scale of a few seconds or minutes, which create a musical phrase or a movement. We believe similar links exist in Finance, as studying markets means both dealing with structural events at the scale of the day, month, or year and also studying events that happen every millisecond in high-frequency trading. State-of-the-art models in this domain use more complex structures [1] than the ones we develop here, but in this approach, we design simpler Deep Learning approaches that are similar to recent ones [2].

The originality of our work comes from the fact that we use a particularly impressive dataset (which we will describe later) that was released in 2019. One of the purposes of this work was also self-learning, which is why the code base is mostly implemented from scratch. One of the main challenges of this problem was to handle the data and go from MIDI files to numerical data that could be used as input in a Neural Network, and that are good representations of the information carried out by music notes.

II.PROBLEM STATEMENT

The exploration of AI in music generation is expanding, yet existing literature reviews are narrow, focusing on limited comparisons and relying on unstructured, selective examinations.

This research aims to fill the gap by systematically reviewing AI-based music generation to understand its scope, trends, and methodologies comprehensively. Notable applications of AI in music, such as the completion of Beethoven's Tenth Symphony, highlight its potential, but the techniques, datasets, and architectures used remain under-explored. This study will investigate the geographical distribution of research, the AI techniques employed, their application in musical composition, integration with music production workflows, and the incorporation of emotion in music. By providing a broad analysis of AI's role in music generation, this research seeks to illuminate the field's current state, identify research needs, and underscore AI's transformative potential in music composition and production.

III.METHODOLOGY

1. Choice of Music Representation:
The study employs ABC notation for training the music generation model, chosen for its simplicity and effectiveness. ABC notation, an ASCII musical format, consists of two main parts: metadata and the musical tune. The metadata includes essential details like tune index, type, time signature, and default note length. The musical content is represented through a sequence of characters, facilitating the training process for polyphonic note sequences via a multi-layered Long Short-Term Memory (LSTM) network.

2. Dataset Preparation and Processing:
Training is conducted on a polyphony ABC notation dataset featuring compositions for piano. The dataset amalgamates tunes a diverse collection of musical styles. Musical notes are represented using integer encoding from a dictionary of 89 unique characters, with the dataset being one-hot encoded for categorization. Training is executed in batches, with each batch comprising 16 sequences, each 64 characters in length.

3. Designing the LSTM Architecture:
LSTM networks, designed to address the challenges of training RNNs on sequential data, are the backbone of the model. An LSTM unit consists of a cell and three gates (input, forget, and output) that regulate data flow, storage, and output across the sequence. This setup allows the network to retain long-term dependencies, crucial for capturing the structural patterns in music. The model incorporates multiple LSTM layers, enhanced with dropout layers to mitigate overfitting and time-distributed dense layers for processing outputs at each timestep. The multi-class classification nature of the problem necessitates a SoftMax classifier, which uses cross-entropy loss to produce normalized class scores. The Adam optimizer is selected for its efficiency in handling sequential input data.

4. Generating Music:
Upon training completion, the LSTM network is primed for generating new musical patterns. The model predicts the next character in a sequence by generating 89 probability values corresponding to the unique characters in the input, selecting the next character based on these probabilities. This process is iteratively repeated, with each new character fed back into the model, culminating in the generation of a new music piece.

IV.LITERATURE REVIEW

Music generation has emerged as a focal point of inquiry within the computational research domain, yielding a plethora of methodologies for algorithmic composition. These methodologies bifurcate into Traditional and Autonomous paradigms. Traditional paradigms operationalize deterministic algorithms predicated on predefined functional specifications to synthesize musical outputs, while Autonomous paradigms leverage iterative learning from pre-existing notational sequences to engender novel compositions.

Initial ventures encompassed algebraic formulations reliant on tree structures for syntactical regulation, and Markovian stochastic processes for model instantiation. After a paradigmatic shift towards artificial intelligence, a variety of models including probabilistic frameworks utilizing Recurrent Neural Network (RNN) variants such as Character-level RNNs and Anticipation RNNs were documented. Additionally, Generative Adversarial Networks (GANs), comprising a discriminator and a generator neural network operating in concert to adjudicate the verisimilitude of generated musical data vis-à-vis authentic compositions, have been actively deployed.

Comparative analyses underscore the proficiency of Long Short-Term Memory (LSTM) networks over GANs in discerning and perpetuating intricate sequential patterns, thereby facilitating escape from repetitive note sequences. Conversely, GANs demonstrate an enhanced capability for rapid conceptual uptake albeit at a superficial level, alongside advantages in expedited training cycles.

V.SCOPES AND LIMITATIONS

This project aims to leverage AI for music generation with potential commercial applications, mindful of legal and technical constraints. The approach involves using copyright-free or public-domain music for training to navigate copyright issues. The project's scope is constrained by computational resources and time, acknowledging that the most advanced generative models require extensive data and computational power, which this setup lacks. The study will focus on evaluating and comparing different models, specifically transformer and LSTM models, acknowledging the limitations this brings to the breadth of possible AI modeling techniques explored.

VI.LIMITATIONS

Legal Constraints: The selection of training data is limited by copyright laws, necessitating the use of public domain or copyright-free music.
Computational Resources: Available computing power and time constraints limit the scale of data processing and model training, affecting the potential complexity and efficacy of the generated models.
Scope of Models Investigated: The study is confined to examining transformer and LSTM models due to the vast landscape of generative AI modeling.
Unexplored Applications: Potential uses of AI in music, such as aiding in composition or completing works by deceased composers, fall outside the scope of this thesis.

VII.CONCLUTIONS

The fusion of computer music technology with deep learning has led to the development of an efficient automatic music generation system based on the LSTM model. The system demonstrates commendable success in generating music autonomously but also reveals areas for improvement, notably in the nuanced representation of musical features. The transformation of music from MIDI format to ABC notation has been identified as a bottleneck, causing a loss of certain melodic elements and resulting in a degree of compactness in the generated compositions. To overcome these challenges, further research is warranted to refine the process of extracting more comprehensive musical information, ensuring a richer and more accurate representation of melodies in future iterations of the music generation system.

VIII.REFERENCES

https://www.researchgate.net/publication/351708912_Music_Generation_using_Deep_Learning

https://www.sciencedirect.com/science/article/pii/S0957417422013537

https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1204/reports/custom/report11.pdf

https://medium.com/@alexissa122/generating-original-classical-music-with-an-lstm-neural-network-and-attention-abf03f9ddcb4

https://blog.paperspace.com/music-generation-with-lstms/

https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5

https://www.midiworld.com/mozart.htm

https://keras.io/