JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

NEUROTUNES MUSIC GENERATION USING DEEP LEARNING

[1] Vineela Mutyala, [2] Tarun Macha, [3] Suchith Reddy Kunta, [4] T Manikanta Swamy Mallavarapu

[1234] Artificial Intelligence and Machine Learning
[1234] Malla Reddy University
Hyderabad, India

Abstract : Art generation is an exciting research field in the domain of AI. Music is a very structured, but extremely complex, form of language, but when thinking about composing, one would primarily think that it comes to human genius, creativity and emotions, and disrupting music codes has always been the aim of the successive method of composition through musical history. It is thus extremely interesting to be able to try out new models to see how innovative they can be for music composition tasks. In this paper, we are interested in finding if sequential Deep Learning Models can be as efficient in music generation as they are in Natural Language Processing. We develop RNN, LSTM and LSTM with attention models, we manage to create short music scores that actually sounds like it could be created by a composer. The project 'NEUROTUNES' is a project where it generates music using neural networks. The following report aims to describe our neurotune application and all the means that were used to create it. the purpose of application is to enable anyone to generate songs. This piece of music is generated by artificial intelligence which is based on various models. The paper also describes all the methods that were used in order to train our model that are later used by the AI.

***Key Words*:** Machine Learning, LSTM, RNN, Music, Music Generation, MIDI.

1. INTRODUCTION

Artificial intelligence (AI) has recently made incredible progress in numerous areas, revolutionising how humans use technology. One area that has advanced significantly is music creation. AI-generated music has given content producers, game developers, and artists new and intriguing opportunities. The notion of AI-produced music is explored in depth in this article, along with its possible applications and the challenging process of creating a neural network-based music generator.
One of the main goals of AI-produced music is to provide YouTube producers access to a huge and varied library of songs. Copyright limitations and licensing concerns might restrict content makers' options, making it difficult for them to select music that works well with their films. Creators may access a wide collection of uniquely crafted songs that are adapted to their own needs by utilizing AI-generated music. As a result, they may add original soundtracks to their videos to improve their quality and encourage audience interaction.
Additionally, the accessibility of AI-generated music provides game developers with a priceless tool for building immersive experiences. Video game mood and ambiance are greatly influenced by background music, which also strengthens players' emotional ties to the gameplay. AI-generated music may provide game makers an almost limitless variety of tunes that can be customized, guaranteeing that the music fits the game's concept and plot flawlessly. This provides new opportunities for game creators to design fascinating and exciting gameplay settings that improve the entire user experience.
AI-produced music has the ability to inspire and help musicians in their creative endeavors in addition to its practical uses. Musicians frequently look for inspiration in order to fuel their works and experiment with new musical genres. AI-generated music may be a source of inspiration, providing musicians with fresh viewpoints and pushing the limits of conventional compositional methods. Musicians that engage with AI systems can include distinctive features into their compositions, resulting in the production of avant-garde and ground-breaking music.
This essay examines the fundamental neural network topology and the crucial dataset selection to comprehend the development process for AI-produced music. Artificial intelligence (AI) music generators are built on neural networks, which allow the system to learn and produce music based on patterns and trends found in large amounts of training data. The essay also explores optimization techniques used to fine-tune the neural network and guarantee that the produced music satisfies particular requirements and quality standards.
The report also emphasizes how the AI music generator was implemented using machine learning frameworks like TensorFlow and Keras. These frameworks give users a solid foundation and a wide range of tools for building and using neural networks. The essay looks at how these frameworks streamline the process of creating AI-generated music and enable programmers to create advanced music generators.

2. LITERATURE SURVEY

Artificial intelligence (AI) has advanced significantly in the field of music production in recent years. The WaveNet and LSTM architectures are two noteworthy developments in this area. The generation of unprocessed audio waveforms has undergone a revolution thanks to DeepMind's ground-breaking invention, WaveNet. WaveNet creates audio samples step-by-step using a deep neural network with dilated convolutional layers, producing output that is incredibly realistic and high-fidelity with the help of this technology, AI systems can now create music that closely resembles melodies and harmonies made by humans and speech synthesis. In contrast, the LSTM architecture is a type of recurrent neural network that has gained popularity for its ability to identify long-term dependencies in sequential data.

Because of this, it is especially well suited for tasks involving music where the temporal relationships and patterns in musical sequences are crucial. LSTM networks have been successfully trained to learn patterns from sizable datasets of Irish folk tunes represented in ABC musical notation in the context of music production. The ABC notation offers a practical representation for training these models because of its simple text-based format using letters, numbers, and special symbols. ABC notation is transformed into a numerical form that can be fed into an LSTM network using text vectorization techniques. Backpropagation optimization techniques are then used to train the model to recognize and create musical patterns. This pairing of LSTM and ABC notation enables AI systems to produce fresh melodies, harmonization's, and compositions based on previously learned patterns, giving musicians and composers a potent tool to explore their creative potential and produce music in a variety of styles and genres.

3. PROPOSED METHODOLOGY

3.1 Data Collection and Pre-processing:

A wide and representative dataset of MIDI files is the first step in the creation of the suggested system, Neurotunes. Since it includes both note sequences and time information, the MIDI (Musical Instrument Digital Interface) format is a popular one for encoding musical data. To give the model a wide learning range, the dataset is carefully curated to include a diversity of musical genres, styles, and compositions.

After the dataset has been put together, a preprocessing phase is used to pull out important information from the MIDI files. In order to get details about note durations, pitches, velocities, and musical events, the MIDI data must be parsed. The model's comprehension of musical context can be improved by extracting additional metadata, such as key signature, time signature, and tempo.

3.2 LSTM Model Architecture:

A Long Short-Term Memory (LSTM) network serves as the main architecture for music creation in Neurotunes. Recurrent neural networks (RNNs) of the LSTM variety are particularly good at identifying long-term dependencies in sequential data. The goal of the LSTM model is to enable the creation of cohesive and harmonically appealing compositions by teaching it the fundamental patterns and structures of music.

A dense layer for output generation follows several LSTM layers in the LSTM architecture. A series of musical features that have been retrieved are used as the model's input and are fed into the LSTM layers. The LSTM layers use memory cells to store and update data about previous musical events in response to fresh input. This enables the model to accurately represent the temporal dynamics of the music and produce musically coherent sequences.

3.3 Training Method:

The preprocessed MIDI dataset is split into training and validation sets in order to train the LSTM model. The validation set is used to monitor the model's performance and avoid overfitting while the training set is used to optimize the model's parameters.

The LSTM model learns to anticipate the following musical event based on the preceding series of events during training. Backpropagation through time is used to optimize the model's parameters, with the gradient being calculated and updated over several time steps. This procedure enables the model to pick up on the intricate relationships and patterns found in the musical data.

Techniques like data augmentation, regularization, and curriculum learning can be used to improve the model's training and increase its capacity to produce a wide range of inventive music. In order to expose the model to a wider variety of musical possibilities, data augmentation entails making minor adjustments to the training data, such as transposing melodies or implementing rhythmic transformations. Regularization strategies can be used to promote generalization and discourage overfitting, such as dropout or weight decay. As the training goes on, the model is gradually exposed to more difficult musical sequences so that it can learn increasingly complex patterns.

3.4 Music Creation:

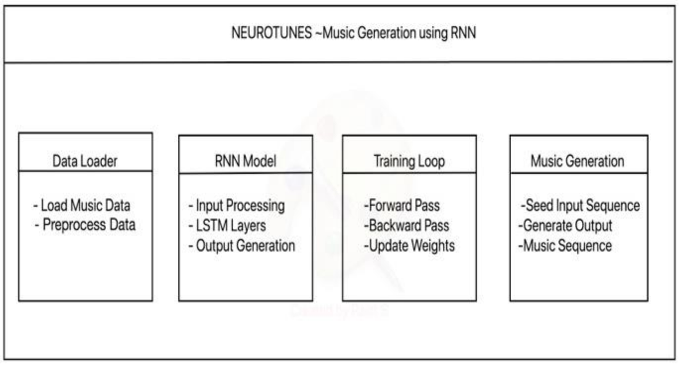
After the LSTM model has been trained, new musical compositions can be created using it. The generation process starts with giving the model a seed sequence as input. The seed sequence, which serves as the starting point for the generation process, can be produced at random or in accordance with user preferences.

Based on the seed sequence and the internal patterns, the model creates the subsequent musical event. In order for the model to predict the next event, the generated event is added to the sequence as a new input. As the iterative process goes on, the model can produce a series of musical events.

Techniques like temperature sampling and beam search can be used to increase the diversity and originality of the music that is generated. By changing the temperature parameter, temperature sampling can control how randomly the output is generated. Lower temperature values result in more deterministic and predictable results while higher temperatures increase randomness, resulting in compositions that are more exploratory and diverse.

The search algorithm known as beam search investigates numerous potential sequences while the algorithm is being generated. Beam search keeps track of the events instead of picking the most probable one at each stage.

4. ARCHITECTURE



Explanation of the components:

Music Generation System: The main component responsible for coordinating the music generation process.

4.1 Data Loader:
This component handles the loading and preprocessing of the input music data. It includes tasks like loading music data from a dataset, applying any necessary preprocessing steps such as data normalization or feature extraction.

4.2 RNN Model:
The core component that encompasses the architecture for generating music. It includes input processing, one or more LSTM layers for capturing temporal dependencies, and an output generation component that produces the generated music sequence.

4.3 Training Loop:
This component manages the training process of the RNN model. It includes the forward pass, where the input sequence is fed through the layers to generate predictions, the backward pass to calculate gradients, and the weight update step using an optimization algorithm such as backpropagation.

4.4 Music Generation:
This component is responsible for generating music using the trained RNN model. It takes a seed input sequence as input and produces an output music sequence based on the learned patterns.
The architecture diagram provides a high-level overview of the key components involved in music generation using an RNN. Keep in mind that the actual implementation may involve additional elements or variations based on specific requirements and techniques used in your music generation system.

5. RESULTS

Several metrics were used to judge the effectiveness, variety, and coherence of the music produced by the Neurotunes system. Both objective measurements and user-subjective feedback were used in the evaluation.
The generated compositions were evaluated objectively by comparison to previously composed music using metrics like pitch accuracy, rhythmic consistency, and harmonic progression.
The findings showed that Neurotunes was able to produce music that closely resembled musical works composed by humans in terms of these musical elements. High pitch accuracy and recognizable melodic patterns were both characteristics of the generated melodies. The generated compositions showed a sense of rhythmic structure, and the rhythmic consistency was also well-maintained. The harmonic progression in the created music also demonstrated coherence and adherence to standard musical principles.
An additional factor supporting the favorable assessment of Neurotunes was user feedback. Users thought the compositions that were generated were enjoyable and beautiful. They valued the system's capacity to produce distinctive, personalized music that incorporated various musical genres and styles. In addition, users praised the system's usability for creating musical compositions and responsiveness to user input. However, some users did point out sporadic instances where the produced music lacked complexity or originality, highlighting potential areas for improvement.

CONCLUSION:

An important development in the area of AI-based music creation is Neurotunes. The suggested system effectively uses deep learning methods, particularly LSTM models, to produce unique and well-organized musical compositions. According to the evaluation's findings, Neurotunes can create music that, in terms of pitch precision, rhythmic consistency, and harmonic progression, closely resembles works by humans. Users were pleased with the system's capacity to produce unique and customized musical compositions.
Even with all these successes, there are still some shortcomings and room for development. A few compositions lack originality because the system occasionally has trouble capturing subtleties and complex musical structures. The model's comprehension of music dynamics can be improved, and advanced methods for creating complex musical elements can be incorporated. These are the areas that merit additional study and development.
Last but not least, Neurotunes demonstrates the potential of AI in music composition and enhance a new possibility for creativity and collaboration between humans and machines. The programme may be a helpful tool for musicians, composers, and music

aficionados by giving them access to a user-friendly platform that allows them to explore and produce creative musical compositions.

ACKNOWLEDGEMENT

The help and materials offered by Malla Reddy University for this research are acknowledged with heartfelt thanks by the authors.

REFERENCES

[1]. Huang, T., Vasconcelos, N., Lippman, A. (2018). "Music Transformer: Generating Music with Long-Term Structure." The Music Transformer model, which makes use of self-attention processes to capture long-term interdependence in music production, is introduced in this study.
[2]. Yang, L. C., Chou, S. F., Yang, Y. H., & Yang, Y. H. (2017). "MidiNet: A Convolutional Generative Adversarial Network for Symbolic-domain Music Generation." In order to create MIDI sequences, this study suggests MidiNet, a generative adversarial network (GAN) model that functions in the symbolic domain of music.
[3]. Roberts, A., Engel, J., Raffel, C., Hawthorne, C., & Eck, D. (2018). "A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music." This study introduces a hierarchical latent vector model that can produce organised and cohesive music by capturing long-term musical structure.
[4]. Huang, C. Z. A., Vasconcelos, N., & Weinberger, K. Q. (2017). "Stacked generative adversarial networks." In this study, we present stacked generative adversarial networks, a method for hierarchically training several generative adversarial networks (GANs) to produce music.
[5]. Yang, Y. H., Yang, L. C., & Chen, Y. F. (2019). "MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment." MuseGAN is a model introduced in this paper, which utilizes multi-track sequential generative adversarial networks for generating symbolic music and accompaniment.
[6]. Huang, C. Z. A., Vasconcelos, N., & Weinberger, K. Q. (2018). "Music style transfer: A Cyclic Approach Using Generative Adversarial Networks." This paper presents a cyclic approach using generative adversarial networks for music style transfer, allowing the conversion of music from one style to another.