Music Beat Generation through AI

Anchal.T, Ayush.S, Piyush G, Sunil K.B, Pushkal.K.S

Department of Information and Technology

Inderprastha Engineering College, Ghaziabad, Uttar Pradesh, India

***Abstract*— Music Plays a significant role in our mental health, but composing a song is difficult for composers. While listening to music, Humans can quickly detect and move to beats. Similarly to Language, Music can be represented as a sequence of discrete symbols that form a hierarchical syntax, with notes being roughly like characters and motifs of notes like a word. Several Theoretical and algorithmic approaches have been proposed to account for this ability. Related to, but different from the issue of how we perceive a beat is a question of how we learn to generate and hold a beat. In this paper, we introduce a Music Beat Generator that will interpolate music rhythms and notes that can be used by music artists, Using Artificial intelligence and Machine Learning. In the context of this project, rhythms are one long measure sequence of notes and rest occurring on natural pulse subdivisions of a beat.**

Music generation using machine learning techniques has been a topic of interest for the past two decades. Music proves to be a different challenge compared to images, among three main dimensions: Firstly, music is temporal, with a hierarchical structure with dependencies across time. Secondly, music consists of interdependent instruments that unfold across time. Thirdly, music is grouped into chords, arpeggios, and melodies – hence each time -step may have multiple outputs.

Keywords – Machine Learning; Artificial Intelligence; Music Beats; Algorithm; Chords; Interpolate; Rhythms.

I. INTRODUCTION

During this time, Computer technology was developed at a swift phase, and this development brought a turning point in music creation. As we all know, composers musician faces great difficulty in creating or composing a good quality sound or music as it requires a lot of time and instruments. In this paper, we proposed our project that enables users to create their music by controlling the music with high-level control parameters. This project is based on the supervised machine learning approach to evaluate music to take the subjective tastes of users into account.

Now we will discuss the material and methods used in our project; we are using machine learning(deep learning), Artificial Intelligence google magenta, and python language for coding. Let us briefly talk about this method and the material used in our projects.

We all know that music compositions require deep music knowledge that cannot be easily created. This project will make music creation easy for those who do not have enough time to learn music deeply. There were many attempts to compose music automatically. The two approaches most general approaches of these attempts are 1). Rule-Based Automatic Composition with musical grammar and 2) Procedural composition with mathematical models such as fractal, cellular automata, or L-system. Therefore, high-level control over the automated composition of various music has not been successfully achieved.

Our project lets the user make their type of music using different instruments soothing to the ears after listening.

We aim to provide an affordable and straightforward solution that can make great beats and rhythms with few clicks and no technical prerequisites.

The module will be both easy to use for the clients and innovative. It would use machine learning and deep learning algorithms to make cohesive and harmonic scores by exploring and mixing various musical ideas. The model would allow us to choose and include multiple instruments and take our input in real-time.

II. MATERIALS AND METHODS

Our project enables users to create music using software that reduces the overall cost of developing the song. Users are only required to know basic things about music. Music is essentially composed of notes and chords. Let us explain these terms from the perspective of any instrument:

A. Note

The sound produced by a single key is called a note.

B. Chords

The sound produced by two or more keys simultaneously is called a chord. Generally, most chords contain at least three essential sounds.

C. Octave

A repeated pattern is called an octave.

D. Beat and Rhythm

A beat is a steady and regular pulse we hear in music, just like an average heartbeat should be regular. The beat and music always is equally spaced and usually does not change rhythm is the actual flow of music through time just as cardiac arrhythmia is irregular rhythm does not have to be equally spaced rhythm usually is built upon the beat, which is why we often see drummers or conductors counting off the beat before the music begins.

E. Sound

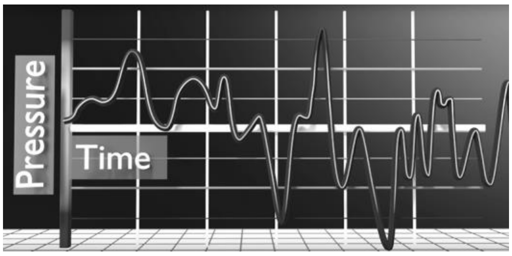

Fig.1 Sound Graph

The vertical axis is pressure or the amount of force the particles bounce into one another. The horizontal axis is time. These high points in the waveform are called compressions, and they are where the particles are compressed and bounce around a lot and thus have a higher pressure. The low points in the waveform are called rarefactions, and they are where the air particles are more spread out and not bouncing around as much and thus have lower pressure. The graph help to visualize sound, which, as we know, is invisible to our eyes. The measurements are taken at a single point in space; in this case, the pressure was measured by a microphone in order for our speakers to duplicate this sound, diaphragm in our speakers have to move similarly. The entire sound takes one-tenth of a second for our speaker to duplicate.

F. Melody

Melody is a series of individual notes when you listen to a melody the melody will be the most obvious layer. It's often higher in pitch or louder in dynamics than the other layers in the background.

Melody can be played or sung on its own and it also can be supported or accompanied by a group of notes. Simple it is where there's music there's melody can't have one without the other.

[2] It is already discovered that what appeals most to people as melody is a fully spun out tune right and that when they get instead an incomplete tune or a theme they begin to have trouble so you can imagine that when they hear music made out of melodies that are even shorter than themes they have even more trouble.

For example that famous opening of Beethoven's fifth again that's so short it's not really even a theme but what is called a motive now a motive can be as little as two notes or three or four.

G. Harmony

Over time, composers started to experiment with what they could do with melody and also started to explore if they could extend what they were doing in music beyond the bounds of pure melody. When they started doing this, they started adding other notes to the mixture as well as the notes that make up the melody. When we combine notes or pitches in music we come up with something called harmony and harmony is probably an important part of music because it allows us to take something that's as simple as a melody and give it a huge amount of extra depth and color just by adding some extra notes.

Google Magenta Library is a Google Brain project that explores the role of machine learning that we can use for creative and artistic purposes. The Google team created an open-source music generating library using machine learning models. The suits include four tools: Continue, Generate, Interpolate, and GrooVAE. Music artists can use the models on their Musical Instrument Digital Interface clips. Currently, the set of tools is provided as a standalone library and as a plugin for the software music sequencer, which was initially designed as an instrument for live performances but is now also being widely used for music generation, recording, arranging, mixing and mastering.

Magenta Studio can also be a valuable tool for musicians looking to break through creative blocks. The Continue tool uses the predictive power of recurrent neural network(RNN)

To generate up to 32 measures of other music following the original drum beat or melody inputs. Moreover, the process not only performs simple predictions of the melody; it creates variations on themes.

H. Artificial Intelligence

Automatically creating music is particularly hard for many reasons. The most hampering being that a simple 3 -minute song a band of humans can quickly memorize has just way too many variables for a computer. The ideal way to train an AI(aka loss function) to become a musician is unknown. The goal itself for us developers is also far from obvious. Are we trying to generate music from thin air or based on some form of input? Or are we trying to generate a system that can accompany a human musician while playing? While we believe currently, there is no reason to panic about their job prospects for musicians. We will look at three companies trying to automatically generate music and see if the first grammy will possibly be given to a data scientist.

The role of artificial intelligence in music beat generators is to apply deep learning networks, a type of artificial intelligence that relies on analyzing large amounts of data. People have to feed software tons of source material covering various sorts of music. The software then analyses data to find patterns, like chords, tempo, length, and how notes relate to one another. By learning from all the input, it can write its melodies.

The majority of music industry Artificial Intelligence use-cases appear to fall into three major categories:

- Music Composition: Companies use AI to create, enhance, and complement music content.
- Music Streaming: Companies are using machine learning and deep learning to recommend.

The Approaches for our project Music Beat Generator are:

Fig.2 Model playing with consistent style throughout

- Collecting Data Sets.
- Selecting the correct models.
- Train models with different algorithms.
- Evaluating resulting models.
- Repeating the above steps with datasets belonging to different instruments.

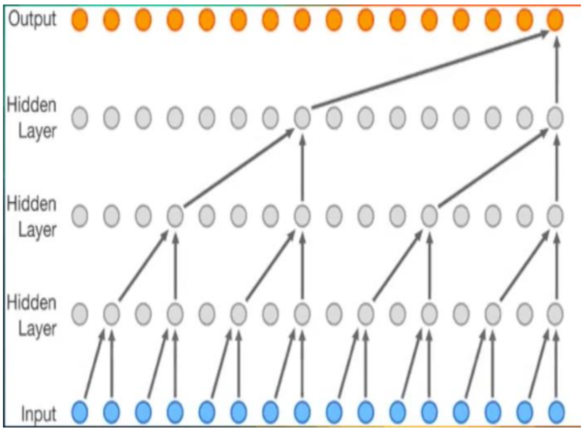

Fig.3 Neural Network to generate the wave

- Music Monetization: Artificial Intelligence platforms are helping artists monetize their music content and generate revenue.

I. *Machine Learning*

In the early 1950s, Iannis Xenakis used the concepts of Statistics and probability to compose music – popularly known as Stochastic Music. He defined music as a sequence of elements (or sounds) that occur by chance. Hence, he formulated it using stochastic theory. His selection of elements was strictly dependent on mathematical concepts. Recently, Deep Learning architectures have become state of the art for automatic music generation. This research paper will discuss two different approaches for automatic music composition using WaveNet and LSTM(Long Short Term Memory) architectures.

WaveNet is a deep learning-based generation model for raw audio created by Google DeepMind. The main objective of WaveNet is to generate new samples from the original data distribution. Hence, it is known as a Generative Model. WaveNet is like a language model from NLP. Given a sequence of words in a language model, the model tries to predict the next word. Like a language model, in WaveNet, given a sequence of samples, it tries to predict the following sample.

[8] Long Short-Term Memory Model, popularly known as LSTM, is a variant of recurrent Neural Networks (RNNs) capable of capturing the long-term dependencies in the input sequence. LSTM has a huge range of applications in different kind of modeling tasks like Speech Recognition, Text Summarization, Video Classification, and so on.

III. IMPLEMENTATION: MUSIC BEAT GENERATOR USING PYTHON:

Magenta is distributed as an open-source python library empowered by TensorFlow. The Library includes utilities for manipulating source data(primarily music and images), using this data to train machine learning models, and finally generating new content from these models.

Traditionally, audio generation models are concatenative. That means generating speech from some text sample, and it utilizes a massive database of speech fragments by picking a few and combining them to create a complete sentence. Same for music. This works, but it is hard to incorporate things like emotion and natural sound when the output is created by putting a bunch of static fragments together. Ideally, we could have an audio generation that is parametric, where all the info required to generate audio is stored in the parameters of the model that we give it. That is what our algorithm is capable of. Instead of generating an audio signal by passing its output through signal processing algorithms, it directly models the raw waveform of the audio signal. Researchers do not do that, but we did. The model we used was a convolutional neural network where each layer had dilation factors that let its interconnectedness grow exponentially the more profound the data flowed through the model. Each generated sample at each step was fed back into the network to generate the next step.

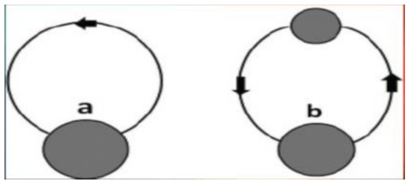

Fig.5 Process keeps repeating

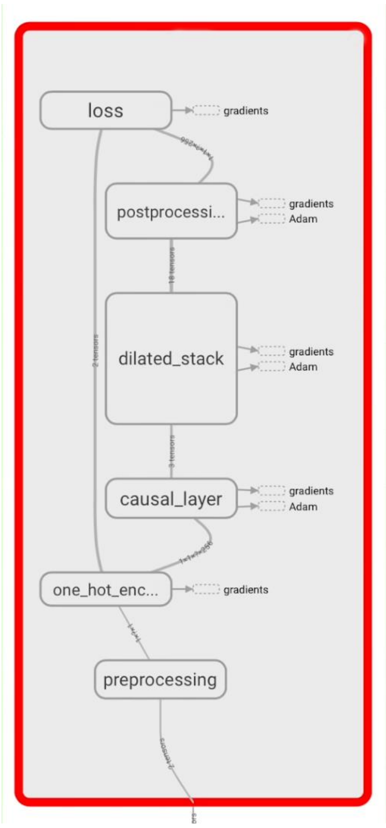

Fig.4 Computation graph for the model

Fig.2 Model forgot about primer almost instantly.

Let us take a look at the computation graph for the model(Fig 4). A single node's input data starts as a raw audio wave. We first formulate the wave so that it is better suited for processing, then we encode it to produce a tensor with many samples and many channels. We feed that into the first layer of the convolutional network, which reduces the number of channels for easier processing, then through a stack of layers with dilated convolution. We combine the outputs of all the layers and increase the dimensionality to the original number of channels. Feed it all to a loss function, which measures how well our training is going, and, finally, the output is fed back into the network to generate the wave of the next time step(Fig-3). This process keeps repeating to generate more and more audio.(Fig-5)

IV. RESULTS

Machine learning is very popular these days, and with open source frameworks and libraries like TensorFlow, developers have access to a wide range of APIs for using machine learning concepts in their projects. Magenta Library, a python library built by the Google team, makes it easier to process and create music and image data in particular.

In this paper, we proposed an automated music beat generation and music evaluation. The composition of music is based on randomness. However, the automated composition with pure randomness seldom produces reasonable or plausible music pieces, so a method restricts the possible output space to be limited within the plausible set of music pieces. In order to achieve such a goal, we employed the concept of mode in the generation process.

The automated composition of music requires two different conflicting capabilities:1) the generation of a restricted set of plausible music pieces and 2)the generation of a variety of music pieces. While the mode is employed to restrict the generation, we also employed density, tempo, and timbre to increase the variety of the generated music pieces. The controlling parameters are converted into binary code for a learning system based on backpropagation. The machine learning-enabled users automatically evaluate the music pieces generated by our system.

The evaluation system made classifies the colossal amount of automatically generated music pieces. The automated evaluation will be successfully utilized to integrate our system to any evolutionary approach to automatically generate desired music or discover the unknown music fit to user-defined objectives.

Our project generates good music according to the user's requirements. However, we had an issue with our model shown in the following figure.

To solve this issue, we divided the data frames; that is, we reduced the audio length of sample data such that only similar motifs were present in a single data frame.

LITERATURE REVIEW

The automation of music composition has been implemented in several ways over the past decade. [3] Some of the oldest methods used were rule-based systems(grammars) and algorithms that utilize randomness like the 'riffology' algorithm presented by Langston(1998). One popular form of grammar used for music was L-Systems. These String -rewriting grammars could produce relatively simple note sequences(1996). In more recent times, the focus for music composition and music generation has shifted towards machine learning and artificial intelligence techniques. [5] This is due to the capability of NNs and deep learning to produce less predictable and more melodically complex melodies than their random algorithm and grammar systems counterparts.

[5] Since there is no way to feed the network with midi values directly and generate music, it becomes essential to encode the music data into a format that is usable by the network.

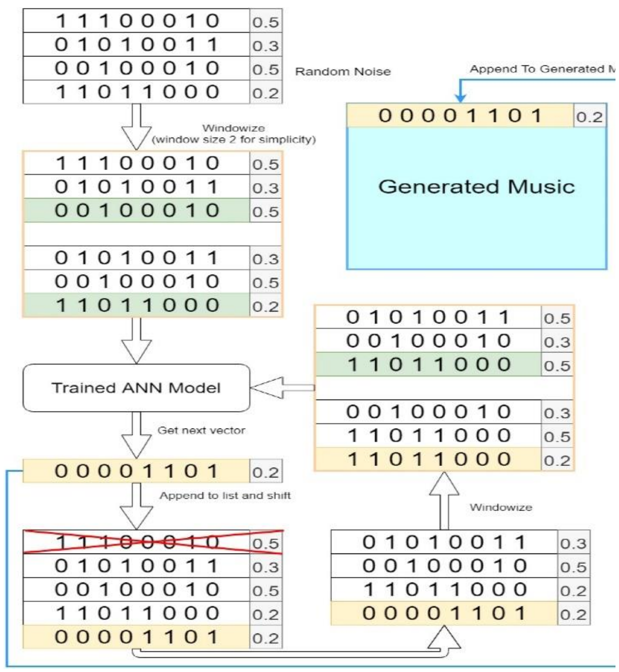

Fig.6 Generation of music after training

At fixed intervals during the network training, the network generates music and then saves it in MIDI format to listen to it and analyze it. The above figure(fig.6) shows how the music gets generated after training. [4] To generate music, The network is fed with a matrix of either noise or a random matrix from the dataset and asked to predict the following vector. After predicting the next vector, we append this vector to the matrix and the generated vector list and pop the last vector from the matrix. Similarly, the same process is performed to predict the following vector. This is repeated many times until a piece of the required length is obtained.

FUTURE WORK

After achieving music rhythm interpolation to an adequate level, next step would be to create a Graphical User Interface for the user. The interface will allow users or operators to provide inputs, data for the music beats generation. The Graphical User Interface will take parameters like bass, treble, pitch, genre, etc. as input from the user. It will be then passed on to the model. The music beats generator will work in the backend of the system and for the frontend, a website will be created. Music Beats Generator can be used in creating background music for video games, stock music, electronic digital music and can be used to generate new rhythm ideas. Beginners can use this tool to create music in no time. Multiple tools can be made available in software for music mixing and matching directly online on the website.

Whether drums, synthesizers, or traditional instruments: Music Beats Generator can be a fully equipped music studio with everything you need to create your sound. Drag and Drop features can be introduced in the system. User can record their own music tracks for which the beats needed to be generated. Music effects can be added with the music track.

CONCLUSION

[1] MusicVAE allows us to create a system which can interpolate two different music rhythms into one. This shows that Artificial Intelligence music mixing and matching is possible and can be used to generate music.

Automatic generation of music allows beginner as well as expert artists to save time on creating music tones. Using this system, they can create, manipulate and synthesize music for their songs. This system is not only limited to music artists. It can be used in making video game music, movie intro music, sound effects, etc. Music Beats Generator shows a lot of confidence that Artificial Intelligence, Deep Learning and Machine Learning can be used in Music Making as well.

REFERENCES

[1] M. Dinculescu, J. Engel and A. Roberts, "MidiMe: Personalizing a MusicVAE model with user data," *Google Research*, 2022. [Online]. Available: https://research.google/pubs/pub48628/.

[2] S. Oore, I. Simon, S. Dieleman and D. Eck, "Learning to Create Piano Performances," *Google Research*, 2022. [Online]. Available: https://research.google/pubs/pub46748/.

[3] A. Bose, Á. Byrne and J. Rinzel, "A neuromechanistic model for a rhythmic beat generation," 2019. [Online]. Available: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006450#abstract0.

[4] D. Horowitz, "Generating Rhythms with Genetic Algorithms," *Aaai.org*.[Online].Available: https://www.aaai.org/Papers/AAAI/1994/AAAI94-253.pdf.

[5] "Generating Piano Music with Transformer", *Magenta*, 2022. [Online]. Available: https://magenta.tensorflow.org/piano-transformer.

[6] C. Huang, A. Vaswani, J. Uszkoreit, N. Shazier, I. Simon, C. Hawthorne, A. Dai, M. Hoffman, M. Dinculescu and D. Eck, "Music Transformer," arXiv.org, 2022. [Online]. Available: https://arxiv.org/abs/1809.04281. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[7] A. Huang, M. Dinculescu, A. Vaswani and D. Eck, "Visualizing Music Self-Attention," *Google Research*, 2022. [Online]. Available: https://research.google/pubs/pub48631/.

[8] "Generation of music pieces using machine learning: long short-term memory neural networks approach," *Taylor & Francis*, 2022. [Online]. Available: https://www.researchgate.net/publication/334999618_Generation_of_music_pieces_using_machine_learning_long_short-term_memory_neural_networks_approach