

Music Generation Using Machine Learning

Jyothin Movva, Satyavarsan ks, Devaraj Acharya

Aziz Nagar, Moinabad Road, Hyderabad – 500075, Telangana, India.

2010030071@klh.edu.in

2010030151@klh.edu.in

2010030040@klh.edu.in

Abstract— The project deals with generation music using inputted dataset. The functionality can be accessed by website which is being hosted by Django framework The application used LSTM and neural networks to train the model and generate the midi file using the model trained.

Keywords—LSTM, NN, Music Generation

I. INTRODUCTION

The days of debating whether artificial intelligence (AI) will have an influence on the music industry are long gone. Artificial intelligence is already being employed in a variety of applications. Now it's time to think about how it will impact the way we create and listen to music. AI automates services, identifies patterns and insights in massive data sets, and helps generate efficiency in the music business, just like it does in other sectors. Companies in the music industry must acknowledge and prepare for the impact of ai on their business; those that do not will be left behind.

Automatic Music Generation is a process of composing a short piece of music with minimum human intervention. The project is about the same topic. Music generation is very important and it can be used in many applications.

This is achieved by recombination of musical phrases extracted from existing music, either live or pre-recorded. Music Generation a viable tool that can and is being used by producers to help in the creative process.

Music generation is one of the interesting applications of machine learning. Music itself is sequential data, it can be modelled using a sequential machine learning model such as the recurrent neural network. This modelling can help in learning the music sequence and generating the sequence of music data. In this Project, we are going to show how we can use neural networks, specifically RNN for automatic music generation.

II. LITERATURE SURVEY

S.no	Authors	Title	Publishing year
1	Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, Yi-Hsuan Yang	MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment	2017
2	Sageev Oore Ian Simon Sander Dieleman Douglas Eck Karen Simonyan	This Time with Feeling: Learning Expressive Musical Performance	2018
3	Nabil Hewahi ,Salman AlSaigal & Sulaiman AlJanahi	Generation of music pieces using machine learning: long short-term memory neural networks approach	2019
4	Prafulla Dhariwal, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, Ilya Sutskever	Jukebox: A Generative Model for Music	2020
5	Li-Chia Yang, Szu-Yu Chou, Yi-Hsuan Yang	midinet: a convolutional generative adversarial network for symbolic-domain music generation	2017

A. MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment:

MuseGAN is a project on music generation. In a nutshell, we aim to generate polyphonic music of multiple tracks (instruments). The proposed models can generate music either from scratch, or by accompanying a track given a priori by the user.

Pros:

An easy logical solution to generate the music. Using the chords changes to train the model is an effective way to train the model to generate new music

Cons:

The dataset uses only midi files which can be less rich compared to music format such as .wav

B. This Time with Feeling: Learning Expressive Musical Performance

Pros:

Uses more musical convention to train the modal instead of chord changes.

Cons:

With a tempo of 120 bmp each beat lasts for 500ms which corresponds to the change of 25ms. This will eventually add up to be a greater change.

C. Generation of music pieces using machine learning: long short-term memory neural networks approach

Pros:

Easy to implement as the tools used are easy to understand. Uses a neural network to generate the music from the given dataset

Cons:

Having many LSTM layers makes the learning progress slower and less accurate than having one or two latest.

D. Jukebox: A Generative Model for Music

Pros:

Music quality is comparatively better than other models. Each of these models has 72 layers of factorized self-attention on a context of 8192 codes.

Cons:

Suffer from hierarchy collapse due to use of successive encoders coupled with autoregressive decoders

E. MIDINET: a convolutional generative adversarial network for symbolic-domain music generation

Pros:

MidiNet performs comparably with MelodyRNN models in being realistic and pleasant to listen to, yet MidiNet's melodies are reported to be much more interesting.

Cons:

The dataset uses only midi files which can be less rich compared to music format such as .wav

Dataset	Characterist	Technique
Midimelody files for theme song	60 midi files	LSTM
Midi piano files	92 midi files	LSTM
Lakh pianoroll dataset	174,154 multitrack piano rolls	GAN
Lo-Fi Hip Hop MIDIs	93 midi files	CNN and GAN
Multi-modal MIREX Emotion Dataset	193 midi files	

(Dataset table)

III. REQUIREMENTS

The required python packages:

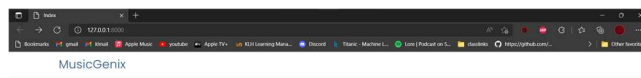
- Pandas – csv file handling
- Keras – NN and LSTM
- Music21 – midi file management
- Pickle – serialize and deserializing

IV. WORKFLOW

We first, found the dataset that we seem fit to use for model training. Once the datasets are collected, preprocessing steps are taken to arrange the midi data set for model training. Model was trained by using LSTM and NN functions from Keras. Multiple models are trained reducing the loss ratio one iteration at a time. Once the ratio reached a desirable, the model generated that iteration was used to generate a new midi file using the help keras for LSTM and NN, Music21 for segmenting and arranging the notes for the midi music file.

V. IMPLEMENTATION

The application is hosted on a web interface. This is achieved by using the Django framework. A single webpage was designed just to input the model for the generation of the midi music file.



Music generation using machine learning

Select a file:
Choose File No file chosen
Upload

(Website hosted using Django framework)

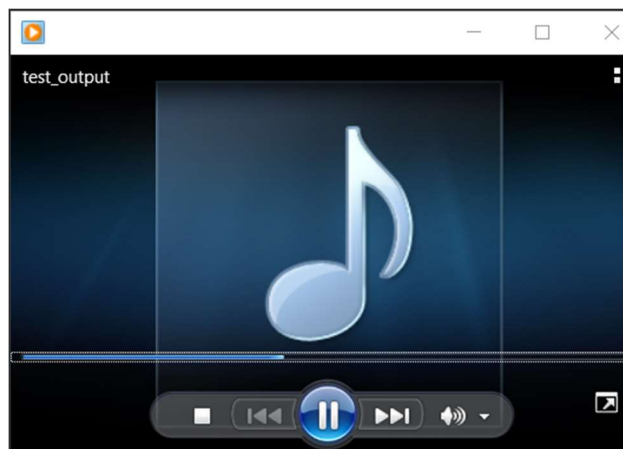
Once the model is inputted on the website, it gets saved in the project root directory and then the python script for generating the output is run which takes the model input from the root directory.

Once the predicting and generating is completed. The output is saved in the project root directory.

manage.py	3/19/2022 09:31	Python Source File	1 KB
test_output.mid	3/19/2022 11:12	MIDI Sequence	6 KB

(Output file generated)

The output is in midi format. To run the midi file, we can use any desirable sound font to play the file.



(Playable music file generated)

VI. CONCLUSION AND FUTURE WORK

The project is fully functional and successfully generates a new music file Django or hosting it in a web server for testing will be appropriate representation of the intended real-life use

Web-Interface can be changed/ modified to enhance the user-experience. Additional options can be added like “genre” to get the output based on chosen genre.

VII. REFERENCES

- [1] *JamBot* – Gino Brunner, Yuyi Wang, Roger Wattenhofer and Jonas Wiesendanger*
Department of Information Technology and Electrical Engineering ETH Zurich Switzerland
- [2] *Convolutional generative adversarial networks with binary neurons for polyphonic music generation* – Hao-Wen Dong and Yi-Hsuan Yang
Research Center for IT innovation, Academia Sinica, Taipei, Taiwan
- [3] *MuseGan* – Hao-Wen Dong*, 1 Wen-Yi Hsiao*, 1,2 Li-Chia Yang, 1 Yi-Hsuan Yang 1 Research Center for Information Technology Innovation, Academia Sinica, Taipei, Taiwan 2 Department of Computer Science, National Tsing Hua University, Hsinchu, Taiwan
- [4] *Lo-Fi Hip Hop Generation* - Zachary Katsnelson