Music Generation using LSTM

Rahul Manusmare
rahul.m@iitgn.ac.in
Indian Institute of Technology
Gandhinagar, Gujarat

Figure 1: Visualising Sample Music Generated

ABSTRACT

The goal of the project is to generate music from scratch of from any partially given sequence of a music. Existing music generation are generalised in terms of their generated music output. Our approach involves music generation according to a specific genre or also in a generalised form.

KEYWORDS

music generation, LSTM, neural network, music dataset

1 INTRODUCTION

Composing music is an art, and composing one need skills, efforts and time. Everyone is having a different taste in music so one music can be pleasing for one individual but the same might not be that

great to other individual in such a scenario composing music which is pleasing for each person at specific mood is very difficult.

The goal is that a person should be able to generate music according to a genre and mood without any previous knowledge of music and without wasting time in thinking the composition and attempting for a good music.

The approach involves the training of LSTM network to recognise the pattern in the music and then predicting the next note according to there patterns define the generation of music using deep learning.

2 DATA

One of the main concern in training the model was getting the right data. we have used MIDI files. Best option for getting the data in the form of MIDI form was because MIDI files dose not store the actual music but is stores the notes of the music which quiet portable and can be easily interpreted.

MIDI for getting the random music generated we have used a mixed data set which consisted of some Jazz, Blues, Pop, Hip-Hop music files.

These music files of different genre are having single or multiple track of only one instrument which is piano. these tracks are being decomposed into different "Notes" and "Chords".

The data was in the fir of Streams, these streams consisted of different attributes of the music files. so in input stream was imported and then according to the type of parts of the elements of the stream they were categorised into Notes and Chords. where Notes were directly flattened and the Chords were Separated with '.' for the different valued of the pitches present in the chords. these different flattened notes and dot separated pitches and notes were of 358. in the case of the mixed data set.

3 METHODOLOGY

3.1 Input data

The initial Data which consisted of MIDI Files were parsed and searched for different steams of music in it. theses stream were separated according to the Instrument type. different instrument were having different streams then according to the type of the streams, if the stream had multiple instruments then the streams were recursively called so that the notes which are to be parsed are separated.

these separated elements of the music file are the parsed to check the type of the element. If the type of the element is note then the notes are then the pitch present in the note can be easily accessed. then the input element are being check if they are of the type Chord. Basically a chord is combination of different pitches so if the element is s chord then the chord is being decomposed into different pitches and thee pitches are being stored in a file.

So, the after parsing the input we get a file which is having a list of different pitches and notes in it.

3.2 Encoding Data

Now, as we are having our files parsed and got different pitches and notes then. before getting these as input to the model we need to encode the data. as on the set of different pitches and notes each element is given a unique id. Then getting this data grouped into groups of 100, we are storing the 101Th element as our y which we need to predict. getting 100 elements without the first element and including the new predicted value as the l01st element of the sequence we are predicting the 102Nd element and thus the window keep on moving.

This data which is being grouped into groups of 100 are then normalises by dividing the input sequences by the total number of different notes and pitches which we got as an input after parsing the data set.

the every 101Th element what we got after a sequence is being is being encoded in the for of one hot encoding.

3.3 Model

As we are dealing with music files we need to retain the memory so as to predict the values so we are using specifically LSTM network as they are the Long short term memories which will retain the sequence in the memory and on that input they will predict the new values.

WE are having three LSTM layers with ,three dropout layer so that our model doesn't get over fit to the input music files then we are also having the the activation function as soft max. we have also defined the loss of the model as categorical cross entropy. We are

using the optimizer as rmsprop as they are good with the recurrent neural network.

3.4 Generating Notes

After getting the encoded normalised input this input is trained on the model to around 100 epochs where we get significant loss. these weights are being saved and then these weights are being used to predict every 101Th Note. Now appending that node to the sequence and then removing the 1St node from the sequence, we get the new sequence which is then feed ed to the network to get the next node and this keeps on repeating till the required notes are achieved.

Finally the results are again encoded into the form of nodes and chords then these notes and chords are converted into a stream

4 RESULT

Evaluation of generated music didn't have any metric to get the melody of a music. So, the music files were upload to sound cloud and got a descent response from the community. there was also a survey conducted during the poster presentation of the project. according to which 10 samples were collected and which scored as 5/5.

5 CONCLUSION

It is concluded that the music generated by the model are melodious and and also music can be generated from some partially initialize music files.

6 REFRENCES

[1] SigurÃřur SkÃžli, âĂIJHow to Generate Music using a LSTM Neural Network in Keras,âĂÌ Towards Data Science, 07-Dec-2017. [Online]. Available: https://towardsdatascience.com/how-to-generate-music-using-a-lstm-neural-network-in-keras-68786834d4c5. [Accessed: 26-Apr-2019].

[2] G. Sharma, âĂIJMusic Generation Using Deep Learning,âĂÌ Medium, 17-Oct-2018. [Online]. Available: https://medium.com/datadriveninvestor/generation-using-deep-learning-85010fb982e2. [Accessed: 26-Apr-2019].

[3] A. Huang and R. Wu, âĂIJDeep Learning for Music.âĂÌ Stanford.edu, 2015. [Online]. Available: https://cs224d.stanford.edu/reports/. [Accessed: 26-Apr-2019].

[4] âĂIJmusic21: a Toolkit for Computer-Aided Musicology,âĂÌ Mit.edu, 2018. [Online]. Available: https://web.mit.edu/music21/. [Accessed: 26-Apr-2019]. âĂŇ