Anish Shah
Nick Torba

Machine Learning MS#0

Project Idea

We aim to generate music using a Recurrent Neural Network (RNN) Restricted Boltzman Machine(RBM) with LSTM architecture schemas (Long Short Term Memory). The inspiration for this architecture stems from many previous projects that successfully generate original works of music. Music is an interesting and intricate realm to replicate using machine learning as it pushes the boundary between human and computational creativity.

Dataset Description

The dataset will consist of a collection of trap MIDI files which are an aggregation of hexadecimal commands that will determine the appearance of a note, the note itself, the velocity of the note, etc. From our research it can be seen that the position temporally, the separation of a note into pitch steps, the temporal positioning of a note in respect to the surrounding temporal position of the notes, the actual note itself in relation to other surrounding notes, and the beat seem to be all important features that can be extracted from the MIDI files.

Planned code/libraries

- The midi package, mido package, or the TonesInTunes MIDI to excel software can be used in conjunction with a csv reader to read the MIDI files
- Extract key features from whatever format the package places it into and clean it into a workable format.
- Numpy to store the cleaned features.
- Tensorflow with packages for all the layers, optimization methods, RNN, LSTM, and RBM to be able to train the collected MIDI data from the stored numpy.
- Take the trained model and use it to produce music which requires conversion of whatever format python put our “predicted midi” in and turn it back into its hexadecimal byte form (CPickle)

Eval/Baseline

Unlike classification problems, there is no hard mathematical baseline to judge our output. Music cannot be judged by deciding whether or not the numbers have been predicted correctly, but only by human ears judging the pleasure and creative impulses they feel when listening. For that reason, we will continually listen to the output of our program to evaluate possible improvements to the melodies, tones, and other musical characteristics.

Planned work for MS#1

In order to ensure a minimum viable product by December 7th, we will need to have our input data ready to be used by our architecture. To achieve this, we need to clean our dataset and extract the features we have chosen to create our music. Referring back to the dataset description, we must collect a sufficient amount of trap MIDI files to convert into the format which will contain the position, note specified by pitch, relative temporal positioning, relative pitch, and beat (and any other features we may deem more necessary).

Collaboration Plan

Fortunately, our team will consistently be able to work on this project together. Because most work will be done as a team in person, we do not feel it is necessary to pre-determine large splits of responsibilities between the two of us. For most of the project, we plan to rotate between working on code and researching theory. In our rotation, we will improve code quality and better our understanding of theory by building off each others ideas and understanding.

http://www.hexahedria.com/2015/08/03/composing-music-with-recurrent-neural-networks/

chrome-extension://ihgdgpjankaehldoaimdlekdidkjfghe/viewer.html#https://cs224d.stanford.edu/reports/allenh.pdf

chrome-extension://ihgdgpjankaehldoaimdlekdidkjfghe/viewer.html#https://cs224d.stanford.edu/reports/NayebiAran.pdf