

Generating tonal piano music using machine learning

Goal

The goal of my project will be to train a model to generate tonal piano music and store it in MIDI files, so that it can then be played back or converted to a score. I chose this project because one of my passions is composing music and playing the piano, and I've recently become fascinated with music generated by Artificial Intelligence, so I wanted to see if I could train my own model to generate tonal piano music.

Dataset

The dataset I will be using will be the MAESTRO Dataset, containing over 200 hours of piano music in MIDI format, available here: <https://magenta.tensorflow.org/datasets/maestro>. Although this dataset also contains the corresponding audio files for these MIDI files, I have chosen only to use the MIDI files to train my model, since I think that having MIDI files with specific notes and timings will be much easier to preprocess and therefore train a model with than audio files. Also, since my goal is for the model to generate MIDI files, it makes more sense for the model's input to also be MIDI files.

I chose this dataset because it was the largest dataset of piano music in MIDI format that I could find. I might supplement it with additional smaller datasets if necessary.

Preprocessing

To preprocess the MIDI data, I will be using the Mido library for Python, a library that is used for working with MIDI messages, ports and files, and that can be found here: <https://mido.readthedocs.io/en/latest/>.

The information that I will need to extract from these files is the pitches, beginning times, and ending times of each note, since those are the features that describe a piece of music.

Machine Learning Model

The machine learning model I am thinking of using is an LSTM Neural Network with Attention. I think this type of neural network will be best for music generation because the Attention will allow the neural network to understand the context of each note in the music, which is obviously very important when generating music. This decision is based on some research I did and on suggestions from my Deliverable Advisor, but I might decide to use a different model once the course advances further and I gain a deeper understanding of the different types of models.

Evaluation Metric

After doing some research online, I found that objective evaluation metrics for machine learning generated music are very hard to come by. Many researchers use surveys asking how much the generated music sounds like music in order to evaluate their model. For a more objective approach, this paper provides some good ideas, but also warns that each model for generating music should be evaluated differently, due to there being so many different styles of music:

https://musicinformatics.gatech.edu/wp-content_nondefault/uploads/2018/11/postprint.pdf

I think I will probably use a survey to evaluate my model, as the evaluation metrics I found online for evaluating machine learning generated music don't seem to be very representative.

Final Conceptualization

I will integrate my model into a Flask web app where people can click a button and a MIDI file will be generated for them, that they can then download and turn into a score using any scorewriter program, or listen to using any program that can playback MIDI files.