Electronic Music Generation with Magenta

Sybil Bingqing Li | Advisor: Prof. Takunari Miyazaki | Trinity College, 300 Summit St, Hartford, CT 06106

Introduction

This project is an investigation of the possibility of artificial intelligence capable of composing music.
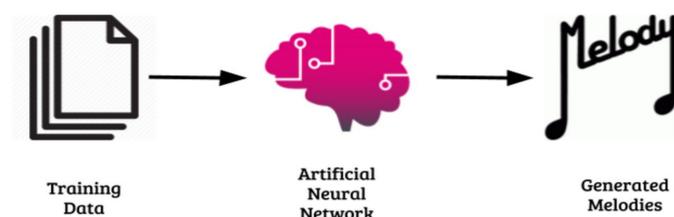
- A recurrent neural network is trained on MIDI files of electronic music to generate melodies based on given input.
- The model achieved 84% accuracy on the training data.
- An interactive demo was built with the trained model, using which a human agent is able to feed input into the model with a MIDI keyboard and get instant feedback from the model.

Motivation

Attempts made to use machine intelligence in creative arenas such as music and art generation are important because they

- Pose a demanding task for the field of computer science
- Challenge of our understanding and assumptions of what intelligence is and what we are capable to create with intelligence

Method

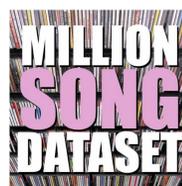Training Data — Artificial Neural Network — Generated Melodies

The idea is to build a recurrent neural network (image below), and training it with MIDI files of electronic music. For faster training, the neural network is trained on NVDIA GPU.

Technologies

TensorFlow   magenta
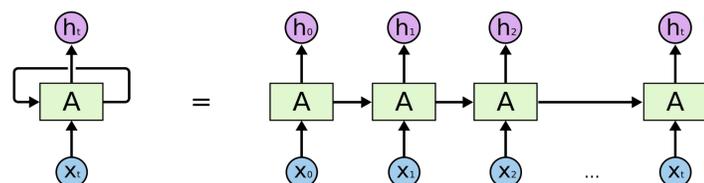
Data Collection

The Lakh MIDI Dataset v0.1

- The Million Song Dataset is a freely-available collection of audio features and metadata for a million contemporary popular music tracks.
- The Lakh MIDI Dataset is a collection of 176,581 unique MIDI files, 45,129 of which have been matched and aligned to entries in the Million Song Dataset.

The million song dataset has genre labeled, thus all MIDI files corresponding to electronic music can be filtered out from the subset of MIDI files from the Lakh MIDI Dataset that has been match to the Million Song Dataset, using the genre label on the Million Song Dataset.

Results and Discussion

Accuracy against Number of Training Steps:
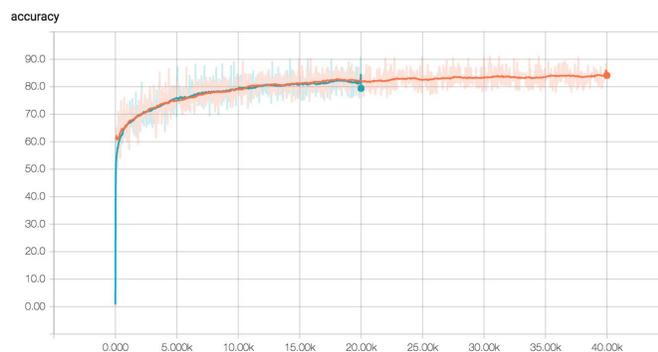
Loss against Number of Training Steps:
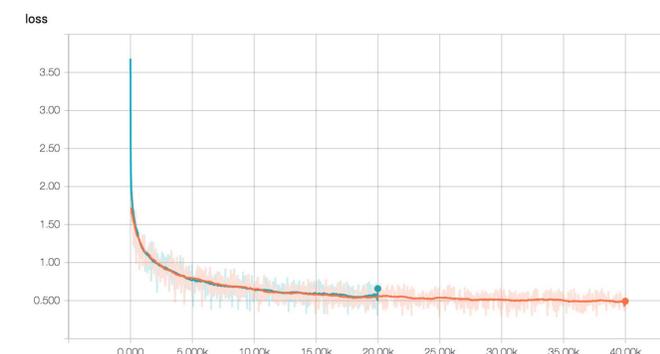
Model Specification: batch_size:64, rnn_layer_sizes:[64,64], num_training_steps=20000/40000

As can be seen from the graphs, the highest accuracy the model is able to achieve is clearly above 80%. An accuracy of 80% shows definitively that the model is not making random guesses, but prediction based on some understanding of electronic music composition.

Interactive Demo

MIDI keyboard — Trained Model — Ableton Live Lite
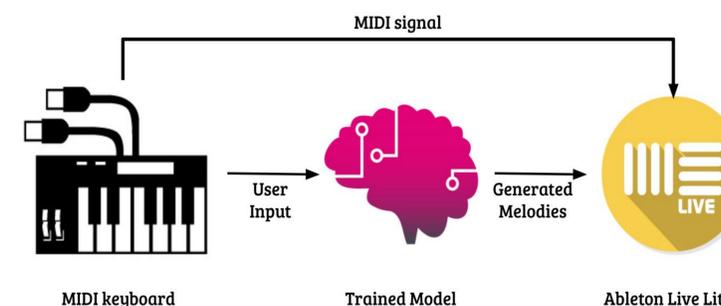
Conclusion

It can be concluded that the results obtained shows promising prospect for an intelligent machine capable of creative tasks such as composing music.

With more training data and computing power, the neural network could have achieved even more impressive results. Though much still needed to be done, expectations remain positive.