Anish Shah
Nicholas Torba

Standard Midi File Classification using a Recurrent Neural Network

Music synthesization has been a domain that has grown with music and the accompanying musicians all throughout time. In our current age, with the ubiquity of technology musicians have has shifted to using synthesizers, in the form of MIDI controllers, and their computers to compose music. Music of all genres has been composed using this methodology, with definable qualities rooted in music theory and recognition of musical patterns by the human listener. Along with this, listeners now have instantaneous access to all this music from a multitude of devices making the need for accessible ways to access a variety music via tags all the more prevalent. As such the task of genre classification is ever growing. To achieve this task an RNN with LSTM was implemented with gradient clipping, dropout, and optimized using ADAM. This was done to take into account the appearance of notes of a song in its temporal context for all instruments. The features were engineered by implementing the PrettyMidi package which converts the hexadecimal instructions of the MIDI file into an array with row instrument, column timestep, and field of note played by an instrument at each time step. For each song their related feature matrix was fed into the RNN which outputted a prediction of the genre of the song. With our RNN we anticipate/have achieved accuracies of 55% without EDM midis. With the presence of EDM midis we anticipate/achieved accuracies of 33%.