Proposal for the Classification of Classical Music Nicholas Angell

This project will require a learning model to analyze a piece of music and determine in which era the piece was composed. These eras would include ​Baroque​ ​(1600–1750), Classical ​(1750–1820), ​Romantic ​(1780–1910), ​20th Century​ ​(1901–2000), and Contemporary ​(1975–present). Including early music periods such as the ​Medieval ​(500-1400) and ​Renaissance ​(1400-1600) eras is an option as well. The model will be reading the music itself instead of any metadata like the composer or year the piece was written. Data will be collected from MIDI files of a specific piece, which holds data about each note in the piece of music.

Fitting an entire piece of music into a single instance would be improbable, so a single instance may include a short segment of the piece, such as two measures worth of notes, three seconds of the middle of the piece, or a collection of thirty or so notes in a row. A learning model should be able to analyze how notes fit in with other notes to create the music. In addition, the specific instruments used in the piece and whether the piece uses an orchestra may help to classify the piece, or you may choose to only select music which utilize the piano or other instruments specifically. Thus, an instance of data may look like this:

Harpsichord? Piano? Violin? Orchestra? Winds? Voice? Note1 Time1 Note2 Time2 ... N Y Y N N N F6 000 A6 000

… Note28 Time28 Note29 Time29 Note30 Time30 Period? . G7 1370 G#7 1450 A 1500 Romantic

Data can be collected from numerous MIDI collections of classical music, such as the following databases:

http://www.kunstderfuge.com https://www.classicalarchives.com/midi.html http://www.piano-midi.de/ https://www.classicalmidi.co.uk

Converting the midi format into a format which can be understood by the model will take some work, but several free tools can be used to assist in this effort, such as midicsv, midiscript, MidiConvert or other github projects. As for classifying the training data, it is generally accepted that each composer fits into one specific era, which can easily be looked up. If there is any ambiguity about the composer’s era, do not use the composer’s music as training data. Select music from several different composers per era so that the model learns the style of the era, rather than the style of a composer.