Milestone 2 Report

Team members: Juliana Choi (88867312), André Santiago Correia (59366682), Harman Sihota (99773566), Charles Wang (87971487)

Main goal of milestone

Decide on the best model for our project and put it to test. Convert data into best format for our model.

Current state of project

For this part of the project, each team member did research in order to find ways to convert our audio files, and to figure out the best model for our solution. Furthermore, we met for the TA office hours in order to ask them for tips on how we should build our model. Below are some of our main sources and best insights taken from each one:

Note Recognition in Python is a medium publication by Ian VonSeggern that guided us through the first steps of processing our audio files. It included important insight that confirmed our suspicions that our dataset would have to be modified (i.e. identify where notes are so that we get short data samples that aren't very populated). Furthermore, it showed code snippets and plots of sample data. Lastly, it provided insight on how to classify the notes. It was one of the most important resources we found.

Similarly, Plotting A Spectrogram Using Python And Matplotlib also provided input on how to turn our audio files into something more manageable and familiar such as images (spectrograms). Meanwhile, Automatic Music Transcription Using Neural Networks is a research paper that was developed with the same prompt as our project, which lead to very important insights, such as "Constant-Q Transform has been applied to transform the audio data to the frequency domain in a logarithmic scale" and "one-hot encoding vectors have been used to represent the output data, i.e., the music notes".

Play Sheet Music with Python, OpenCV, and an Optical Music Recognition Model is an interesting article that describes research done to decode and play sheet music. Despite doing the opposite of what we want to do (we want to have sheet music as the outputs, not inputs), it brings great insight on how we can build our model, as they also used Machine Learning to solve the problem. They used a convolutional neural network combined with a recurrent neural network and trained the model with stochastic gradient descent (SGD), which aimed at minimizing the connectionist temporal classification (CTC) loss function. Our team hopes to further research these terms and to discuss with people who have more experience with ML.

One of the main challenges our project faces is when multiple notes are played at once since our model is more likely to misinterpret those once fed with just single notes. For that challenge, insight was taken from the article Neural networks for musical chords recognition. They built a model entirely with that purpose, so our group will research how we can deal with the problem of wanting a model that does multiple things at once.

Lastly, Note Detection is a project that turns music signs into sheet music using digital signal processing tools in Matlab. This particular piece of research is important and relevant to our project as our final goal is to produce sheet music out of audio files.

Besides the research done with week, some of our members have also started testing resources found online, such as Pydub, which allows you to manipulate audio files by slicing, changing the volume, etc., Librosa, for converting our audio files into spectograms, and Pretty_Midi and Midi2audio, which allowed us to use our MIDI files.

Technical Challenges

Converting spectograms into waveform graphs
Choosing the best way to output predictions for our model so that we can compare with the actual sample's outputs for training purposes
Creating a convolution model from scratch
Adapting our model so that it can recognize single notes and chords/multiple notes at once.

Team tasks

Juliana Choi

- Research of how to turn audio into images and whether that is the best model for our project (DONE)
- Create the notebook file that our team will collectively work on (DONE)
- Write Milestone 2 report (DONE)
- Research on how our model can tackle different aspects at once (IN PROGRESS)

Andre Correia

- Create the layout of our notebook file which will guide us and organize our code (DONE)
- Start coding some cells of our notebook (DONE)
- Continue working on the non-ML component of the project by learning more about HTML, JS and CSS. (IN PROGRESS)

Harman Sihota

- Continue looking for other sources of data and acquiring more data (DONE).
- Convert .ps files into .png (DONE)
- Learn how to convert midi files into possibly more useful formats such as a spectrogram image (DONE)

- Research on methods of comparing sheet music and midi/audio files for training (IN PROGRESS)

Charles Wang

- Research into data acquisition via Youtube (DONE)
- Start preparing some code for model creation (DONE)
- Research on identifying sounds as notes in sheet music for model (IN PROGRESS)