Milestone 4 Report

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

Main goal of milestone
Complete project by finding ideal loss function and model, finish implementing website, and have both the model and website connected.

Current state of project
With guidance from the professor we found that our labels were too complex and would likely lead to poor accuracy. So we decided on a new data representation. We split the input audio into 0.5 second segments (to avoid multiple notes in the same segment) and made spectrograms for each of these. Using the midi files, we made labels for each of these images. There were 257 classes:
- 0 represented no note start or end in the 0.5 second interval
- 1-128 represented midi notes 1-128 starting in the interval
- 129-257 represented midi notes 129-257 ending in the interval

Our model is a classifier that identifies which class an input spectrogram belongs to. We used transfer learning from the Resnet 18 model and adapted it to our needs. The biggest alteration made to the Resnet model was the fully connected layer, which used to have 1000 features but now has 257. Since this was a classification task, we used cross entropy loss.

This model achieved an accuracy of 47% in 15 epochs on a dataset consisting of approximately 10,000 samples.

Furthermore, this week we created the non-ML part of the project, a website that connected directly to our model in order to generate the MIDI file required for the sheet music. Currently, the website is being hosted on the members' computer's local host. The website asks for a Youtube link and once inputted, it uses the library youtube-dl in order to download the audio file. Then, the wav file is input into our model and outputs a MIDI file. Finally, the user can input the MIDI file into MIDIToSheetMusic in order to get the sheet music corresponding to the video.

<u>Team tasks</u>

Juliana Choi
- Write the Final Report
- Film and edit the video
- Provide input on the model and dataset according to music background
- Write the Milestone 4 Report

Andre Correia
- Create the HTML and CSS files for the website (front-end)
- Code the backend of the website
- Connect the model to website's input and debugged it
- Filmed parts of the video

Harman Sihota
- Experiment with different data reps and loss functions
- Change data representation and create the new dataset to improve accuracy
- Decide on new model and loss function according to the new data rep
- Train model and experiment with training hyperparameters to maximize accuracy
- Write code to use the model to convert a users input audio file into a midi file using the trained model

Charles Wang