

Daily Log

Monday November 11

I tried to write the MFCC values of all my Hot 100 songs onto a text file to read into my CNN, but I kept getting a File Not Found error. Eventually I found its source: there's a file in my Hot 100 folder that was saved in a weird file format. I'll have to investigate because I have no idea how that happened while I was downloading the playlist. As far as I know, there's only one file like that in the folder; the rest of the songs' MFCCs were being successfully calculated until I ran into that error.

Tuesday November 12

I need to organize song data in order to feed it into my neural network, so I installed the Pandas Python library. Pandas is used to organize data from file formats such as .xlsx (spreadsheets). Using 5 songs as test data, I recorded a few of their audio features and MFCC values into a spreadsheet to start formatting with Pandas.

Thursday November 14

I found a paper written by a group of researchers in Australia who are using MFCC for speech recognition. I'm reading parts of it right now. I also found their contact information and I'm considering which questions to ask them. I'm thinking I'll ask for a step-by-step explanation of how MFCC is calculated and how to read a MFCC graph.

```

1 #Victoria Agrinya
2 #Last update: 11.12.19
3
4 from __future__ import print_function
5 import keras
6 from keras.models import Sequential
7 from keras.layers import Dense, Dropout, Activation, Flatten
8 from keras.layers import Conv2D, MaxPooling2D
9 import os
10 from keras.datasets import boston_housing as bh
11 import pandas as pd
12
13 songs = pd.read_excel(io="/Users/vicki/Documents/Senior_Research/Song_Feature_Data.xlsx")
14

```

Figure 1: Uploading data from an Excel document using the Pandas library.

Timeline

Date	Goal	Met
10/28/19- 11/1/19	Continue work on convolutional neural network and download Hot 100 songs	Yes
11/04/19- 11/08/19	Continue work on CNN	Yes
11/11/19- 11/15/19	Continue work on CNN	Yes
11/18/19- 11/22/19	Continue work on CNN and email MFCC researcher	In progress
11/25/19- 11/29/19	Continue work on CNN	In progress

Winter Goal

Run CNN and have it predict Spotify song popularity score (on a scale of 0-100) with at least 80 percent accuracy.

Reflection

I can't start inputting data into my CNN until I get the data formatted into arrays, and I need to have each category of data (MFCC, tempo, etc.) represented properly in an array for Keras to read it properly. That's where I am with my CNN right now. I'll also be reaching out to different researchers to get a comprehensive "crash course" on how exactly MFCC works so I can cover my bases and be able to explain it in presentations, since it is an important factor in my project. I'm also figuring out how to upload files to GitHub and I'll get that done in the coming week.