Ethics Review Notes Per Sprint

First Sprint (discussion of Project + finding dataset)

- After our first sprint we had discussed that our project is mainly targeted towards music listeners who most likely use other music applications (like youtube, spotify, etc).
- With our current plan users may have their preferences of music tracked (such as saving what music they consider positive or negative). This is only meant to help recommend the user on what they would like and not to be used for any other reason.
- We currently do not have a clear idea on how the dataset we had found had rated each song or if there is any certain pattern on certain genres\artists having a certain positive or negative score which could indicate some biases in the dataset itself.

Second Sprint (starting coding of project such as parsing data + Beginning of Neural Network training)

- From what we had started with this sprint there were not many changes to the overall ideas ethically discussed since the last sprint.
    - Standard user assumed is still the same
    - Keeping track of user preferences are still considered (only to be used to help them)
- We had started training our neural network based on our dataset and based on the spectrograms we created from our database we could try to rate the emotions for another song in the future as well.
- We wished that the songs themselves had more categories in them to help break down the data (like gender of singer or genre of song) and due to that we were not able to check to see if there was some sort of pattern of difference in emotion score between different categories.

Third Sprint (continuing Neural Network training + change on the output of Neural Network)

- With the end of the Neural Network training and change in output (now it sends back an emotion score of positive_ neutral_negative instead of just the highest rated emotion) we can show more information to the user on why we played a song for them (ex: why one song can have a more positive feel to it then another).
- We also decided during this sprint that instead of keeping track of each user and their preference we may just end up adding their preference to our whole training model (if the user chooses to do so) which would reduce surveillance of each individual user themself.
    - This should now make us be more wary of our target audience. If our application attracts a certain group of people with the same opinions of song that could end up skewing our neural network to overly thinking one kind of music is positive or negative which could impact minority users (the ones who do not follow the same sentiment as the large majority of users using our program)
        - Once again as mentioned in an earlier sprint we currently believe that our standard user should be music listeners who use web apps to listen to

their songs. From this we can assume that the users will be those with access of the internet (also could look into demographics of those most likely to use internet and music applications) so we can assume with our most recent change to how our neural network is going to be updated that this group will be the one to skew our AI as mentioned in the previous bullet.

Fourth Sprint (HTML front end + backend)

- Other than adding an interface during this sprint we had also decided on removing the user being able to correct if our AI had made the correct assessment of a song emotion based on the user given song (due to time constraints).
- Due to this option getting removed we are now relying solely on what the AI had been trained on so if the AI happens to be wrong we will no longer be able to correct it and update our neural network to accommodate for it, thus continuing incorrect assessments of songs.
    - Continuous incorrect assessments of our songs not being corrected can create a negative outlook on the songs itself and the singer of the songs.

Model Card - Helping Lions Cry Since 2021

Model Details

- Our AI will give the user a song based on what emotion they click on (positive, neutral, or negative)
- The user will have an opportunity to give a youtube url to have our AI give a feedback on how positive, neutral, and negative it is (given in percentage)
    - (Not quite implemented in the HTML) The user should then be able to say if they believe our AI was correct or not and if not fix the score.

Intended Use

- To allow the user to listen to a song they are in the mood for
- Not intended to give a concrete statement on if a song is positive, neutral, or negative rather a probability that the song could convey that emotion and suggest that song

Factors

- The main factor of our AI were the spectrogram of the song, which shows the frequencies of the song
    - Other potential factors that we could not account for in our project were the genre (certain genres can have a certain frequencies to it) and singer of the song (certain singers can choose to keep themselves to a certain frequency)

Metrics

- The performance metrics of the prediction is based on the training accuracy and validation accuracy. The existing model we trained has a training accuracy of 98.69% with validation accuracy of 80.53%. The performance is satisfactory since we split the songs to ten equal 3 seconds excerpts. The second performance metrics is based on training loss and validation loss. The existing model has training loss of 0.0382 and validation loss 0.8013. (Lower the better, where 0 is no loss) However, the actual prediction accuracy will be based on how the neural prediction model fits into the user's preferences.

Training Data

- The IMAC dataset (https://gaurav22verma.github.io/IMAC_Dataset.html)
    - Over 4000 songs tagged with a pos_neut_neg score (0 to 5 for each score)

Ethical Considerations

- Skewing dataset to overly represent certain songs over others and the potential of falsely representing a song based off user input (can be done in terminal but not html)
- If personal data were to have been collected to personalize experience for users make sure the data remains private and transparency is made between developers and users on what happens with their data

Caveats and Recommendations

- Dataset could not be broken down into things such as genre of song or gender of singer which could be used to indicate any sort of bias in dataset