Jaclyn Welch  
CS 491 Capstone II Computer Science  
Spring 2022  
Option 2 – Audio Surveillance System

Engineering Notebook – Sprint 2

Week 1:

I spent this week looking for new sample data to train our model with. The previous data were different voice commands, which was not as applicable to the goals of our end product. I was looking for distinct sounds that one would want to note in a room without video surveillance that were .wav files and had at least 200 files to divide between the training and testing data. I found a large dataset of “Urban Sounds” that had been used by a different project for classification. It contained 10 different distinct sounds with labels, .wav files, classification data, and had between 400 and 1200 files per sound. I downloaded the sounds and looked at the code this other group used for their classification, but it involved embedding the files with metadata when it’s created and before it’s entered into the model. I saved this code to reference, but will be sticking with our current model.

Week 2:

I filtered through the different classifications of data I found last week and filtered out two relevant sounds that my group decided to use: gunshots and an AC unit. We decided it would be helpful to recognize the AC unit to filter it out of concerning sounds, and a gunshot because it without video surveillance, this would allow us to identify where a gunshot is coming from in a room. This posed a challenge I wasn’t familiar with as all of the data was grouped together, but I was able to find a file manager that could filter using the metadata embedded in the files. At this point, Caleb had finished the visualization software and wanted to join the classification part of the system. For us to work on this together, we decided to run the code in colab instead of on the local system. This was also something completely new – this was actually my first exposure to colab – so we spent a few hours uploading the filtered datasets and adjusting the code in colab. We took some time after this to go back to the test plan and update it with the adjustments we’ve been making to the design and how we’re interacting with and planning to test it.

Week 3:

We started this week by touching up more of the documentation – we’ve made several big adjustments to our design as we’ve been troubleshooting, and our documentation was already missing some core components during the first version submission. I worked on updating mainly the SDD and parts of the SRS, and noting which sections needed diagrams created. After this Caleb and I went back to our colab code, which we were having trouble running with the uploaded data. I did not recognize what the error the model threw meant, which was “Bad audio format for WAV”. I did a deep dive online and could not understand what this error meant. It indicated that the .wav files were encoded as a different number of bits than 16-bit, which is the format that our model was expecting. I continued diving into this subject and could not find how to address this other than when you’re originally creating the .wav file and initially formatting it. This caused a full stop in our process, because we had no idea how to address this issue on the data or on the code side. Our next step we decided would be to look for new data that was already formatted to be 16-bit, although we were not 100% sure how to check this.

Week 4:

We were stumped at how to improve our demo with any new sounds, as our last demo had the sample data of the voice commands, and our new data wasn’t working with the model yet. We continued trying to troubleshoot the data issue, and we found new sample datasets that had more alarm and traffic related sounds. This data did not seem to have the exact same issues as the previous data, but we were still not able to get it up and running for the demo.