Week 7 (October 17-October 21)

Tools used:

Python

Goal:

Objective: Graph histograms for all participants

Establish Density vs category graph (and then apply threshold such as max duration of crying annotation occurrence)

Replace vocalization density with crying density

Statistics to many parameters

Purpose: For Kaya’s grant, we want to be able to claim that we analyzed through a certain amount of generated data, and can accurately say that we predicted 50%+ of crying occurrences. We need statistics to be able to claim this, and we need graph visualizations to be able to back the data up and interpret the data.

Results:

Files Created:

densityAnalyze.py -> created to graph actual/detected vs crying density

**pregenerate()**

**if 1 is selected**

* Must put all csv’s to be analyzed in /categories (In this case, all participants of 1 minute episodes with labeled categories “yes” or no”
* Parses through each csv and compiles them to one csv with all occurrences of “yes” and “no” rows

**If 2 is selected**

* Parses through each of the csv’s and calculates the statistics for both the duration of the entire episode and the duration of the labeled (category) episodes

**Execute()**

* Reads from compile.csv and finds the likelihood of crying by counting how many cry’s are detected and how many of them are labeled yes. Yes/detected will give us an idea of how accurate it is. On top of that, I incremented the densitys by .1 and rounded them to the nearest tenth
* Allows user input for certain constraints (to optimize the accuracy). For now, you can specify a certain threshold (in seconds) of how long the duration should be for
* each crying. The higher it is, the higher the accuracy, and the less occurrence per density.

Methods changed:

**Datavis.py**

plotHistogram

* Now plots the histogram using a logarithmic scale in the x axis. The base for the log is coded to be 2, and the max number of power is determined by the max duration. An xlim (min,max) is used to limit our view to preserve the size of the bins.

Action:

* generated histograms for each participant
* Computed statistics for each participant’s episodes and crying annotation durations
* Graphed the relationship between density and categorization