**Readme:**

The code base for this homework consists of three main python notebooks and a script for extraction of speech features.

1. **classification\_text\_sj2995.ipynb:**

This notebook consists of:

1. the code to extract the LWIC textual features from the csv data provided
2. Statistical analysis between Word count and the action tags (shown using bar chart)
3. Training tabular deep learning model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
4. Training XGBoost model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
5. **classification\_speech\_sj2995.ipynb:**

This notebook consists of:

1. Code to extract speech features from csv containing combined data of speech and textual features
2. Statistical analysis between Average Pitch and the action tags (shown using bar chart). Have also shown the analysis between Speaking Rate and the action tags, however that does not show any prominent trends.
3. Training tabular deep learning model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
4. Training XGBoost model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
5. **classification\_speech+text\_sj2995.ipynb**

This notebook consists of:

1. Code to extract both textual and speech features from csv containing combined data of speech and textual features
2. Training tabular deep learning model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
3. Training XGBoost model on the extracted data. Calculating the accuracy, F-1 scores and have shown the Confusion Matrix
4. **Speech\_featuresextraction\_sj2995.ipynb:**
5. Contains the code to split the audio files based on the start and end times mentioned in the csv
6. Further for each interval obtained, it uses parselmouth to extract the speech based features such as Mean Intensity, Mean Pitch, Jitter, Shimmer etc. (Using the approach used in HW1)
7. Converts the extracted data into a csv file