**Project Progress Documentation**

*Research by Jacqueline Zhu | Supervised by Dr. Shiva Akbari | SleepdB Lab*

# ***7.17.2025***

* Completed batch1\_avg\_0719

# ***7.23.2025***

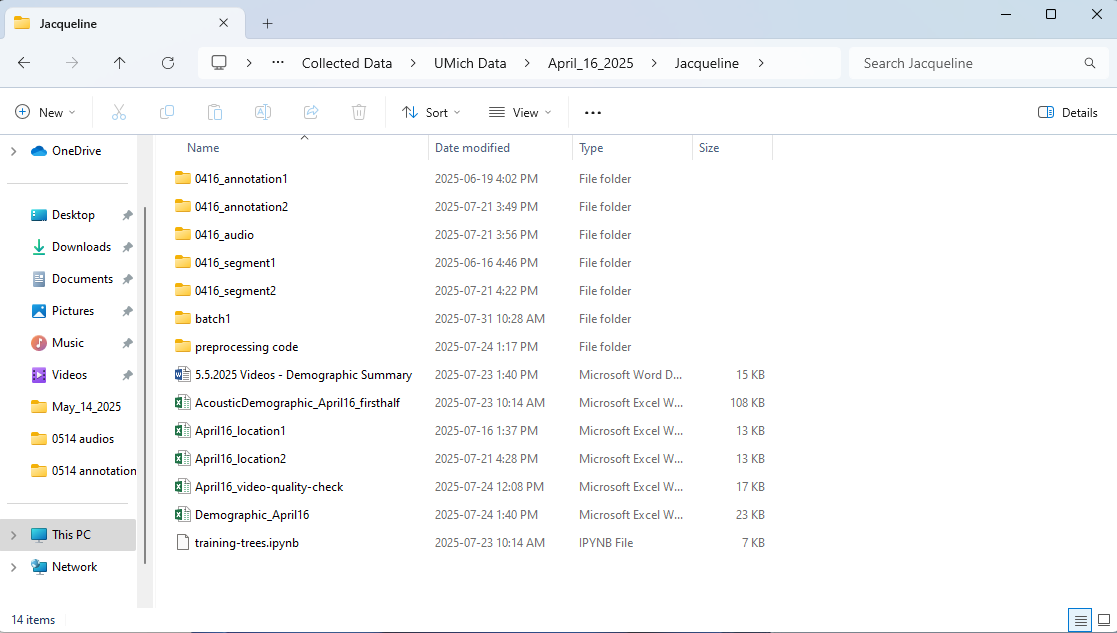
* Brainstormed research update presentation
* Outlined 8 different data split methods:
  + Five vowels together, five vowels separate (see, sahh, so, soo, set), average of five segmented vowels, median of five segmented vowels

# ***7.26.2025***

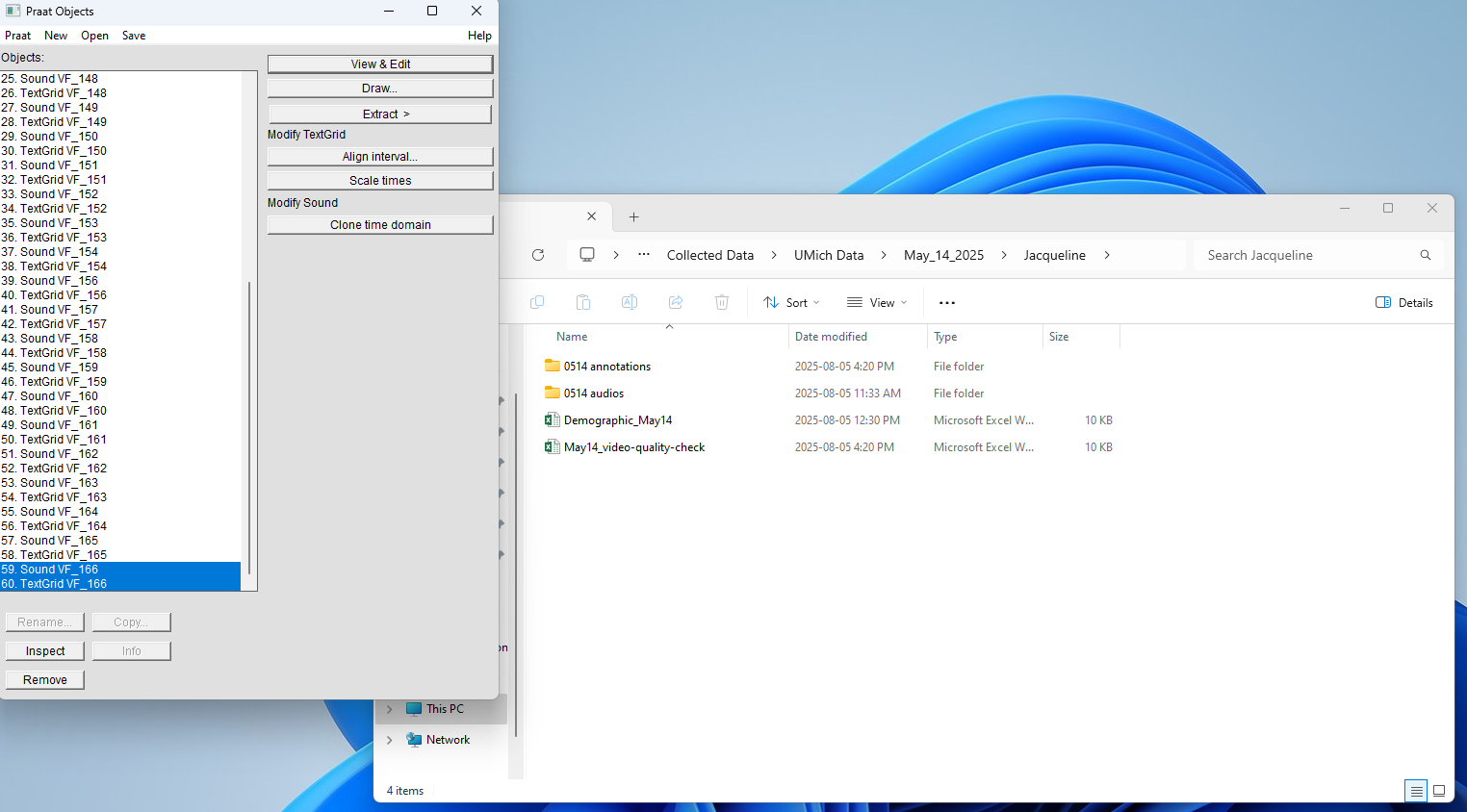
* Started batch2\_all1\_0729

# ***7.29.2025***

* Completed batch2\_all1\_0729
* Prepared research update presentation



# ***8.5.2025***



* Started and finished annotating all the vowels from the May\_14\_2025 folder
* Wrote comments on all the videos and audios in “May14\_video-quality-check.xlsx”

# ***8.6.2025***

* Extracted all the features and AHI information from 5.14.2025 dataset

# ***8.11.2025***

* Reorganize the April16\_May14 datasets – combined all three batches to train on vowels
  + Need to extract the filename + vowels from April 16 \_ batch 2
* Train models on individual vowels
* Split control vs cases for model training

**Table 1: count of the number of samples available per each five vowels**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| see | so | soo | sahh | set |
| 99  01\_DSCF0148.wav | 100 | 100 | 100 | 100 |

* Not all recording samples contain all five vowels | reasons:
  + instructor’s confusion between pronunciation of /so/ vs /soo/ and /see/ vs /set
  + participant’s confusion between pronunciation of /so/ and /soo/

**Table 2: count of the total number of samples available per control and case**

|  |  |
| --- | --- |
| control | case |
| 295 | 204 |

# ***8.12.2025***

* pearson/spearman coefficients, and mutual information on original data
* same statistical significiance methods applied for soo vowel
  + mutual importance high for var
  + need to try other four vowels
* is lost on what to do next because random forest also looks pretty bad lol

# ***8.13.2025***

* MFCC\_0 → Overall log-energy of the frame (roughly total power).
  + MFCC\_0 (-200 to 50): log-energy can vary a lot depending on loudness of the frame. Big negative values correspond to very quiet frames.
* MFCC\_1, MFCC\_2 → Shape of the spectral envelope, low-frequency variations.
  + MFCC\_1 (120 to 260) and MFCC\_2 (-100 to 60): capture large-scale spectral slope differences. The scale depends on how DCT was applied.
* MFCC\_3–MFCC\_12 → Finer spectral details, higher-frequency variations.
  + MFCC\_3–MFCC\_12 (-60 to 40): represent higher-frequency envelope details. These are usually more stable because they capture subtle spectral shape differences rather than overall energy.

Meeting with Shiva:

* Should try MLP (smaller models, less layers)
* Try my method of not log scaling all the features
* Redo batch 2 code to generate similar graphs and get correlation coefficient (R) instead of R\_squared, or fit multiple correlation coefficients onto the graph

# ***8.14.2025***

If the simple, transparent models aren’t capturing the relationships, try a neural net (MLP) with minimal complexity, then increase depth/width as needed.

* Read and learned about MLPs
* 04162025\_05142025
  + Soo\_scaled = only specific features were applied log
  + Soo = no features were transformed / scaled
* 04162025\_05142025 / soo
  + Soo in all features were applied yeo johnson