CogID: A Speech Recognition Program for Early Detection of Alzheimer's Disease

Brianna Broderick, BA, BS; Si Long Tou, BS; Anna Hua, BS

UNIVERSITY OF MICHIGAN

Background

The current process for diagnosing dementia and Alzheimer's Disease (AD) is expensive, complicated and time consuming. Testing requires an individual to be brought into unfamiliar clinical settings to see a variety of specialists which can cause added distress for an individual already experiencing confusion. Repeated visits are often needed and the current lack of geriatric specialists can mean being put on wait lists or long periods between visits. Early detection and diagnosis of AD and other dementias is crucial and can allow for proper management of symptoms.

Our solution is CogID, a tool to detect possible features of AD from an individual's speech.

Data

The data used in this project came from the Pitt Corpus, part of the Dementia Bank[1]. Data selected came from 142 subjects diagnosed with AD and 96 healthy elderly control subjects. Subjects completed multiple visits over time, totalling 234 interviews from participants with AD and 223 from controls.

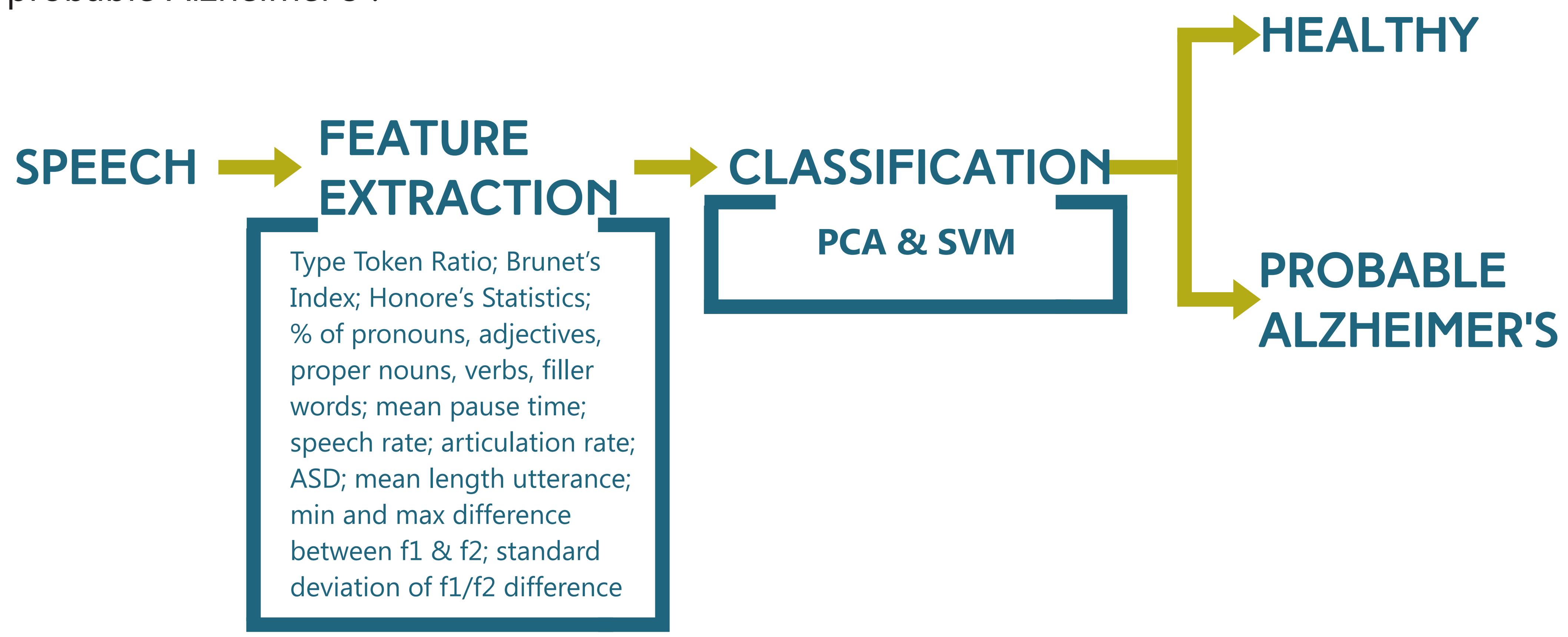
The selected interviews included audio files and transcripts of participants completing the Cookie Theft Picture Description Task from the Boston Diagnostic Aphasia Examination[2]. Prior to classification, all files were processed to exclude any data from the interview facilitator.

Methods

Through literature reviews, the team identified linguistic features that are known and hypothesized to be affected by AD to analyze in the CogID system. Both acoustic (unfilled pauses, speech rate) and semantic (lexical richness, filler words) features were included as components to be analyzed.

Semantic features were extracted using natural language processing and part of speech tagging with the interview transcripts. Acoustic features were extracted from audio files using Praat software[3], which allows for analyzation of speech at the syllable level. These 16 components combine to create a feature vector where each vector is representative of one participant interview.

These feature vectors were used to train and test the CogID classification system. When given a selected participant from the testing group, the system outputs a prediction, either "healthy" or "probable Alzheimer's".

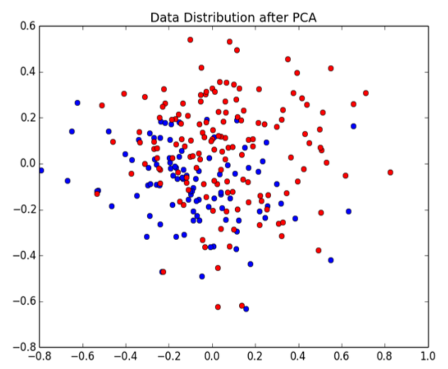


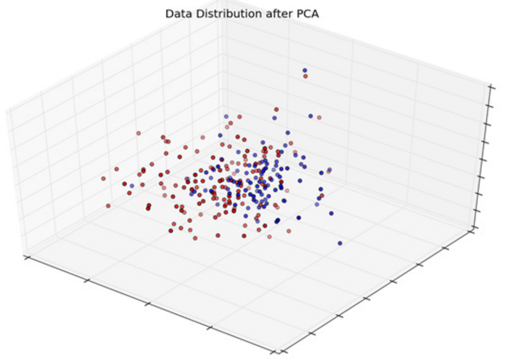
Classification & Results

A variety of classification models were run including K-Nearest-Neighbors (KNN), and Support Vector Machine (SVM) including both linear and radial basis function (RBF) kernels. These models were run using two approaches 1) aggregating all visits for a participant into one feature vector, 2) using only one visit from each participant. Given the small sample size, leave-one-participant-out cross validation was used for testing.

Classification & Results Cont.

To account for the large variation within this data, Principal Component Analysis was applied to reduce the feature vectors to 3 variables, this significantly improved classification performance.





After training and testing multiple models, the team's most successful model used SVM with an RBF kernel and used the combined visit data. The model achieves an accuracy of 0.7327 and a relatively high recall of 0.8296.

Conclusions

This initial prototype has great potential for future use. By adding features and improving the accuracy, the system could have a variety of applications for aiding in the diagnosis of Alzheimer's and other dementias.

Areas for future work:

» Improving the accuracy of the system to be appropriate for use in a clinical setting.
» Expanding the system's capabilities to work in real time.
» Including more robust features including motor coordination and goodness of pronunciation.
» Expanding the system to differentiate between dementia subtypes.
» Training the system longitudinally to recognize progression of dementia stages.

References

1. Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease: description of study cohort and accuracy of diagnosis. Archives of Neurology, 51(6), 585-594.
www.dementia.talkbank.org
2. Goodglass, H., Kaplan, E., & Barresi, B. (2000). Boston Diagnostic Aphasia Examination: (BDAE3). S.l.: Pearson Assessment and Information.
3. Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.40, retrieved 11 May 2018 from http://www.praat.org/