Biomarker-Based Optimization For Alzheimer's Disease Classification

Team 36: **Kerena Belony, Nicholas Chong, Aiyana Cubias, Jordana Kasaato, Amber McBorrough**
Advisor: **Barnaly Rashid, Hristina Schlaggar, David Salat (Niji Neuro Corp)**

There is currently a gap in the clinical diagnosis of Alzheimer's Disease (AD) as patients may have Alzheimer's pathology, but not have a clinical diagnosis of AD. This can lead to misdiagnosis and late diagnosis. Through the use of morphometric features from brain MRI images, we extracted data from features including the entorhinal cortex, hippocampus, and mid-temporal volume to build machine learning models to determine the likelihood that a patient has AD and to predict the likelihood of a patient developing AD. We generated high confidence groups of patients and controls based on their AD biomarker status to enhance these models. Then, we reviewed if the biomarker status was consistent with how individuals were originally classified as patients or controls in the database. To develop a more accurate method of classifying this disease, we separated our data into training and testing sets according to biomarker based confidence levels. The training data (high confidence) was used to build our algorithm, while the testing data (low confidence) was used to test the algorithm. The data were then applied to three different classification learners: Support Vector Machine (SVM), Naive Bayes, and Random Forest. We found that using biomarker confidence decreased the misclassification rate for each classifier. SVM, Naive Bayes, and Random Forest classification models' misclassification rates decreased by 15.8%, 5.6%, and 5.1%, respectively. These results suggest that biomarker-based confidence rankings are important to include in classifiers for AD to provide more efficient support for the clinical diagnoses.